AD-A104 088 TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY F/G 12/1

AN ALTERNATIVE ESTIMATOR FOR THE MAXIMUM LIKELIHOOD ESTIMATOR F--ETC(U)

JUN 81 F SAMEJIMA N00014-77-C-0360

UNCLASSIFIED RR-81-1 NL

1 of 2

AD

A104088

RESEARCH REPORT 81-1

LEVEL II

AD A104088

AN ALTERNATIVE ESTIMATOR FOR THE MAXIMUM LIKELIHOOD ESTIMATOR FOR THE TWO EXTREME RESPONSE PATTERNS

FUMIKO SAMEJIMA

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TENN. 37916

JUNE, 1981

Prepared under the contract number NOOO14-77-C-360,
NR 150-402 with the
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Research Report 81-1
2. GOVT ACCESSION NO.: AD-A104 088
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): An Alternative Estimator For The Maximum Likelihood Estimator For The Two Extreme Response Patterns
5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Dr. Fumiko Samejima
8. CONTRACT OR GRANT NUMBER(s): N00014-77-C-0360
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Psychology, University of Tennessee, Knoxville, Tennessee 37916
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PE: 61153N; PROJ: RR 042-04; TA: RR 042-04-01; WU: NR 150-402
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research (Code 458), Arlington, VA 22217
12. REPORT DATE: 29 June 81
13. NUMBER OF PAGES: 108
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Operating Characteristic Estimation
Tailored Testing
Latent Trait Theory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(Please see reverse side)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

In the methods and approaches we have developed for estimating the operating characteristics of the discrete item responses, the maximum likelihood estimate of the examinee based upon the "Old Test" has an important role. When Old Test does not provide us with a sufficient amount of test information for the upper and lower part of the ability interval, however, it is likely that we obtain response patterns which are either all lowest item scores or all highest item scores, and the resultant maximum likelihood estimates are negative or positive infinity. Although such a test is undesirable for us to use as the Old Test, to a certain extent we can salvage the situation by providing some alternative estimator for these two extreme response patterns. Following RR-80-3, "Is Bayesian Estimation Proper for Estimating the Individual's Ability?", in the present paper, such an estimator is proposed and discussed.

AN ALTERNATIVE ESTIMATOR FOR THE MAXIMUM LIKELIHOOD ESTIMATOR FOR THE TWO EXTREME RESPONSE PATTERNS

ABSTRACT

In the methods and approaches we have developed for estimating the operating characteristics of the discrete item response, the maximum likelihood estimate of the examinee based upon the "Old Test" has an important role. When Old Test does not provide us with a sufficient amount of test information for the upper and lower part of the ability interval, however, it is likely that we obtain response patterns which are either all lowest item scores or all highest item scores, and the resultant maximum likelihood estimates are negative or positive infinity. Although such a test is undesirable for us to use as the Old Test, to a certain extent we can salvage the situation by providing some alternative estimator for these two extreme response patterns. Following RR-80-3, "Is Bayesian Estimation Proper for Estimating the Individual's Ability?", in the present paper, such an estimator is proposed and discussed.

The research was conducted at the principal investigator's laboratory, 409 Austin Peay Hall, Department of Psychology, University of Tennessee. Those who worked in the laboratory and helped the author in various ways for this research include Paul S. Changas, Melanie Perkins, Charles McCarter, Eva Curlee, C. I. Bonnie Chen and William J. Waldron.

TABLE OF CONTENTS

I Introduction

In a previous study (Samejima, RR-80-3), it has been observed that Bayesian estimation has a characteristic which contradicts the principle of objectivity of testing, because of its bias caused by the effect of the prior, in estimating the examinee's ability from his or her performance in the test. Thus a fairly widespread belief among psychologists that Bayesian estimation is better than the maximum likelihood estimation because of the additional information, the prior, and because of the fact that it provides us with finite estimates for all the possible response patterns, should be seriously reconsidered and dismissed. In contrast to the Bayesian estimators, the maximum likelihood estimator has the characteristic of asymptotic unbiasedness, and it has been shown (Samejima, 1977a, 1977b, 1977c, and RR-77-1) that this unbiasedness holds as a good approximation even with a relatively short test and with mediocre values of the test information function for the interval of ability, or latent trait, of our interest.

When the test information function of our test assumes reasonably high values throughout the range of ability of our interest, the probability with which an examinee obtains one of the two extreme response patterns, i.e., the set of the lowest item scores and that of the highest item scores, is negligibly small. In such a case, it is almost certain that the maximum likelihood estimator provides us with finite estimates for all the examinees and will not give us any

inconvenience in either research or practice. The fact of the matter is that we should construct and use such tests for both purposes. In practice, however, very little consideration of such nature and theoretical insight has been taken in constructing tests and the subsequent use of the tests. Many researchers and users of tests casually pick up existing tests and use them for varieties of purposes, and blame the maximum likelihood estimation for the fact that it provides us with negative and positive infinities for some examinees as their ability estimates, and turn to the Bayesian estimation simply because it does not produce infinities. Scientific examination reveals, however, that this is nothing but a disguise; the simple fact is that the test itself fails to have enough power to estimate the examinees' ability levels (cf. Samejima, RR-80-3). We must accept the fact that every test has a finite range of ability for which it can estimate ability levels accurately enough, and avoid the pretense that it can do so outside of that range of ability by the use of such an inadequate information as a prior.

With this basic understanding in mind, a question will arise as to whether there is any way of expanding this range of ability a test has, without turning to any inappropriate information, and without sacrificing our scientific honesty and the objectivity of testing. If this is possible, then it will contribute to our research and practice, since we could use a wider range of existing tests for our purposes with our appropriate selections.

A positive answer to the above question has been given (Samejima, RR-80-3) for the situation in which the number of test items is relatively small, by proposing an alternative pair of estimates for the two extreme response patterns, which will replace negative and positive infinities resulting from the maximum likelihood estimation. The present study is a continuation of the previous study, in which the concept of these alternative estimates is expanded to cover the situation where the test has a larger number of test items.

II Comparison of the New Estimate θ^*_V with Several Other Estimates

Let θ be ability, or latent trait, which assumes any real number, such that

$$(2.1) \qquad -\infty < \theta < \infty .$$

Let g $(=1,2,...,n)$ denote an item, and x_g $(=0,1,2,...,m_g)$ be a graded item response to item g . The operating characteristic, $P_{x_g}(\theta)$, of the graded item response, or item score, x_g is defined as the conditional probability, given ability θ , with which the examinee obtains the item score x_g for item g . In the normal ogive model, this operating characteristic is defined by

$$(2.2) \qquad P_{x_g}(\theta) = (2\pi)^{-1/2} \int_{a_g(\theta-b_{x_g+1})}^{a_g(\theta-b_{x_g})} e^{-u^2/2}\, du ,$$

where a_g (>0) is the item discrimination parameter and b_{x_g} is the item response difficulty parameter which satisfies

$$(2.3) \qquad -\infty = b_0 < b_1 < b_2 < \ldots\ldots < b_{m_g} < b_{(m_g+1)} = \infty .$$

Let V denote the response pattern, or a vector of n item scores such that

$$(2.4) \qquad V' = (x_1, x_2, \ldots, x_g, \ldots, x_n) .$$

By the assumption of local independence (Lord and Novick, 1968), the *operating characteristic of the response pattern*, $P_V(\theta)$, or the conditional probability, given ability θ , with which the examinee obtains the response pattern V , is the simple product of the n operating characteristics of the graded item scores, such that

$$(2.5) \qquad P_V(\theta) = \prod_{x_g \in V} P_{x_g}(\theta) \ .$$

The maximum likelihood estimate, $\hat{\theta}_V$, of ability θ for the examinee whose response pattern is V is obtained by using this operating characteristic $P_V(\theta)$ as the likelihood function $L_V(\theta)$, or, equivalently, as the solution of θ for the equation

$$(2.6) \qquad \sum_{x_g \in V} A_{x_g}(\theta) = 0 \ ,$$

where $A_{x_g}(\theta)$ is the basic function for the item score x_g , which is defined by

$$(2.7) \qquad A_{x_g}(\theta) = \frac{\partial}{\partial\theta} \log P_{x_g}(\theta) \ .$$

The item response information function, $I_{x_g}(\theta)$, for the item score x_g is obtained from the basic function, or directly from the operating characteristic, by

(2.8) $$I_{x_g}(\theta) = -\frac{\partial}{\partial\theta} A_{x_g}(\theta) = -\frac{\partial^2}{\partial\theta^2} \log P_{x_g}(\theta) ,$$

and the item information function, $I_g(\theta)$, is defined as the conditional expectation of the response pattern information function, given θ , such that

(2.9) $$I_g(\theta) = E[I_{x_g}(\theta)|\theta] = \sum_{x_g=0}^{m_g} I_{x_g}(\theta) \; P_{x_g}(\theta) .$$

We can write for the response pattern information function, $I_V(\theta)$, such that

(2.10) $$I_V(\theta) = -\frac{\partial^2}{\partial\theta^2} \log P_V(\theta) = \sum_{x_g \varepsilon V} I_{x_g}(\theta) ,$$

and the test information function, $I(\theta)$, is defined as the conditional expectation of the response pattern information function, given θ , such that

(2.11) $$I(\theta) = \sum_V I_V(\theta) \; P_V(\theta) .$$

It can be shown that the test information function, which is defined by (2.11), is also the sum of the n item information functions, so that we can write

(2.12) $$I(\theta) = \sum_{g=1}^{n} I_g(\theta) .$$

One of the important and useful characteristics of the maximum likelihood estimate is that, asymptotically, it distributes normally with θ and $[I(\theta)]^{-1/2}$ as the two parameters (Samejima, 1975). It has been shown (Samejima, 1975, 1977a, 1977b, 1977c and RR-77-1) that this convergence to the normality of the conditional distribution of the maximum likelihood estimate is fairly fast, and even with a relatively small number of test items and a mediocre amount of test information this asymptotic normality can be used as a good approximation to the conditional distribution of the maximum likelihood estimate, given ability θ, when the operating characteristic of the item score x_g follows the normal ogive model. If the number of items is too small and so is the amount of test information, this approximation will not hold, however.

Let V-min and V-max denote the two extreme response patterns, such that

$$(2.13)\qquad \begin{cases} V\text{-min}' = (0,0,0,\ldots 0) \\ V\text{-max}' = (m_1,m_2,m_3,\ldots,m_g,\ldots,m_n) \;. \end{cases}$$

In such models as the normal ogive model and the logistic model on the graded response level (Samejima, 1969, 1972), the maximum likelihood estimate for the response pattern V-min is negative infinity, and that for V-max is positive infinity. In such a situation as described in the preceding paragraph, the probability with which we obtain negative

or positive infinity as the maximum likelihood estimate is no longer negligibly small. The approximate unbiasedness of the maximum likelihood estimate, therefore, cannot be attained in such a situation. The situation wll be salvaged if we define a pair of estimates, θ^*_{V-min} and θ^*_{V-max}, such that

$$(2.14)\quad \begin{cases} \theta^*_{V-min} = [\frac{1}{2}(\theta_c^2 - \underline{\theta}^2) - \sum\limits_{\substack{V\neq V-min \\ V\neq V-max}} \hat{\theta}_V \int_{\underline{\theta}}^{\theta_c} P_V(\theta)\, d\theta] \\ \qquad [\int_{\underline{\theta}}^{\theta_c} P_{V-min}(\theta)\, d\theta]^{-1} \\ \theta^*_{V-max} = [\frac{1}{2}(\bar{\theta}^2 - \theta_c^2) - \sum\limits_{\substack{V\neq V-min \\ V\neq V-max}} \hat{\theta}_V \int_{\theta_c}^{\bar{\theta}_c} P_V(\theta)\, d\theta] \\ \qquad [\int_{\theta_c}^{\bar{\theta}} P_{V-max}(\theta)\, d\theta]^{-1}, \end{cases}$$

where $\underline{\theta}$ and $\bar{\theta}$ are the lower and upper endpoints of an appropriately defined interval of θ, and θ_c is a critical value of θ below which the operating characteristic $P_{V-max}(\theta)$ assumes negligibly small values and above which $P_{V-min}(\theta)$ assumes negligibly small values, and use them as the substitutes for the negative and positive infinities of the maximum likelihood estimate, respectively (cf. Samejima, RR-80-3).

Since every test has only a finite number of items, and the amount of test information is limited (Samejima, RR-79-1), any test is informative

enough in estimating the examinee's ability only when his or her ability lies within a subset of the entire range of ability. This subset may be a single, finite interval of θ , or a set of several intervals, depending upon the combination of test items and their characteristics. In many cases, however, the test information function of a test of our interest assumes high values only for a single, finite interval of ability θ , and, therefore, the subset is a finite interval, as is the case with LIS-U (Indow and Samejima, 1962, 1966), which was used in the previous study (Samejima, RR-80-3) as an example of a short test. In such a situation, the interval, $(\underline{\theta},\bar{\theta})$, which was introduced in the preceding paragraph, can be considered as the subset. By virtue of the substitute estimates, $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, for the negative and positive infinities of the maximum likelihood estimate, this subset, or interval, has been enlarged, and, moreover, we can obtain an approximately unbiased estimate of ability for this range of ability. We define the estimator θ^*_V such that

$$(2.15) \qquad \theta^*_V \begin{cases} = \theta^*_{V\text{-min}} & \text{for } V = V\text{-min} \\ = \theta^*_{V\text{-max}} & \text{for } V = V\text{-max} \\ = \hat{\theta}_V & \text{otherwise .} \end{cases}$$

Hereafter, we shall call this estimator θ^*_V the <u>modified maximum likelihood estimator</u>.

Table 2-1 presents the discrimination parameter, a_g , and the difficulty parameter, b_g , of each of the seven binary test items of LIS-U , which follows the normal ogive model on the dichotomous

TABLE 2-1

Item Discrimination Parameter, a_g , and Item Difficulty Parameter, b_g , of Each of the Seven Items of LIS-U .

Item g	a_g	b_g
1	1.031	-0.860
2	1.695	-0.520
3	1.020	-0.220
4	0.800	-0.030
5	1.111	0.190
6	1.389	0.470
7	1.370	0.760

response level, and whose item characteristic function is given by (2.2). The test information function, $I(\theta)$, and its square root, of LIS-U are also shown as Figure 2-1, by solid and dotted lines, respectively. Figure 2-2 presents the operating characteristics of the two extreme response patterns, V-min and V-max , of LIS-U by solid and dotted lines, respectively, and the position of θ_c by an arrow. This critical value of θ is defined as the point at which the product of $P_{V\text{-min}}(\theta)$ and $P_{V\text{-max}}(\theta)$ is maximal.† It turned out that $\theta_c = -0.0088$, and $P_{V\text{-min}}(\theta_c) = 0.0027$ and $P_{V\text{-max}}(\theta_c) = 0.0031$. These values satisfy the requirement in defining $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ that for $\theta < \theta_c$ $P_{V\text{-max}}(\theta)$ assumes negligibly small values and so does $P_{V\text{-min}}(\theta)$ for $\theta \geqslant \theta_c$ (cf. Samejima, RR-80-3).

The values of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ have been obtained for eleven different intervals of $(\underline{\theta},\bar{\theta})$ in the previous study (Samejima, RR-80-3), and the regressions of the estimate θ^*_V on ability θ have also been illustrated for these eleven cases. It has been observed that the approximate unbiasedness of the modified maximum likelihood estimate θ^*_V holds better for smaller intervals of $(\underline{\theta},\bar{\theta})$, while the violation of the unbiasedness becomes more conspicuous as the interval becomes larger. Table 2-2 presents the values of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ obtained upon each of the eight smallest intervals of $(\underline{\theta},\bar{\theta})$, with the square root of the test information function at $\theta = \underline{\theta}$ and $\theta = \bar{\theta}$, respectively, together with the upper bound of the discrepancies between the regression

---

†There is a typographical error in RR-80-3, and on page 84, line 2 from bottom, "minimal" should be replaced by "maximal."

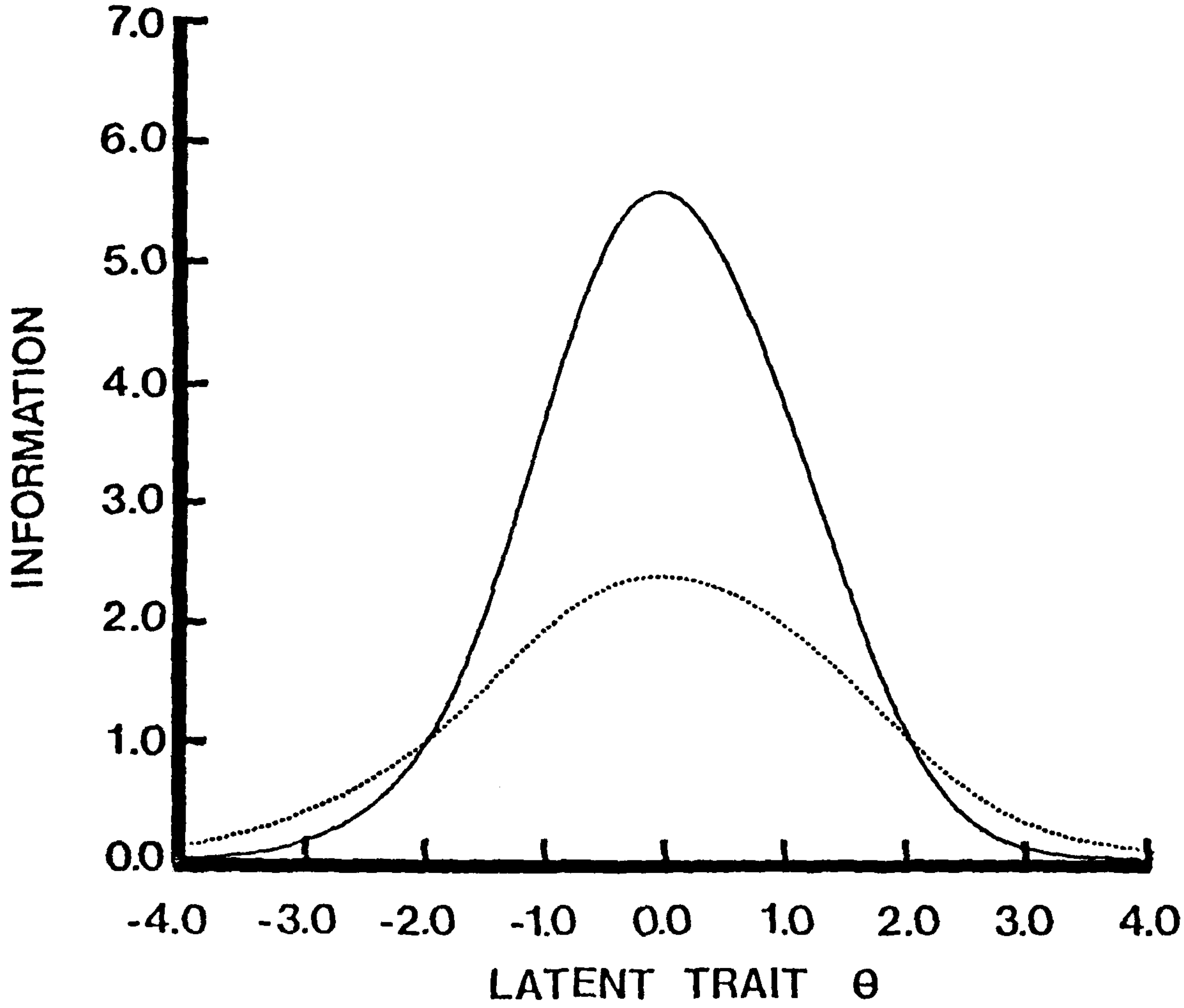


FIGURE 2-1

Test Information Function (Solid Line) and Its Square Root (Dotted Line) of LIS-U.

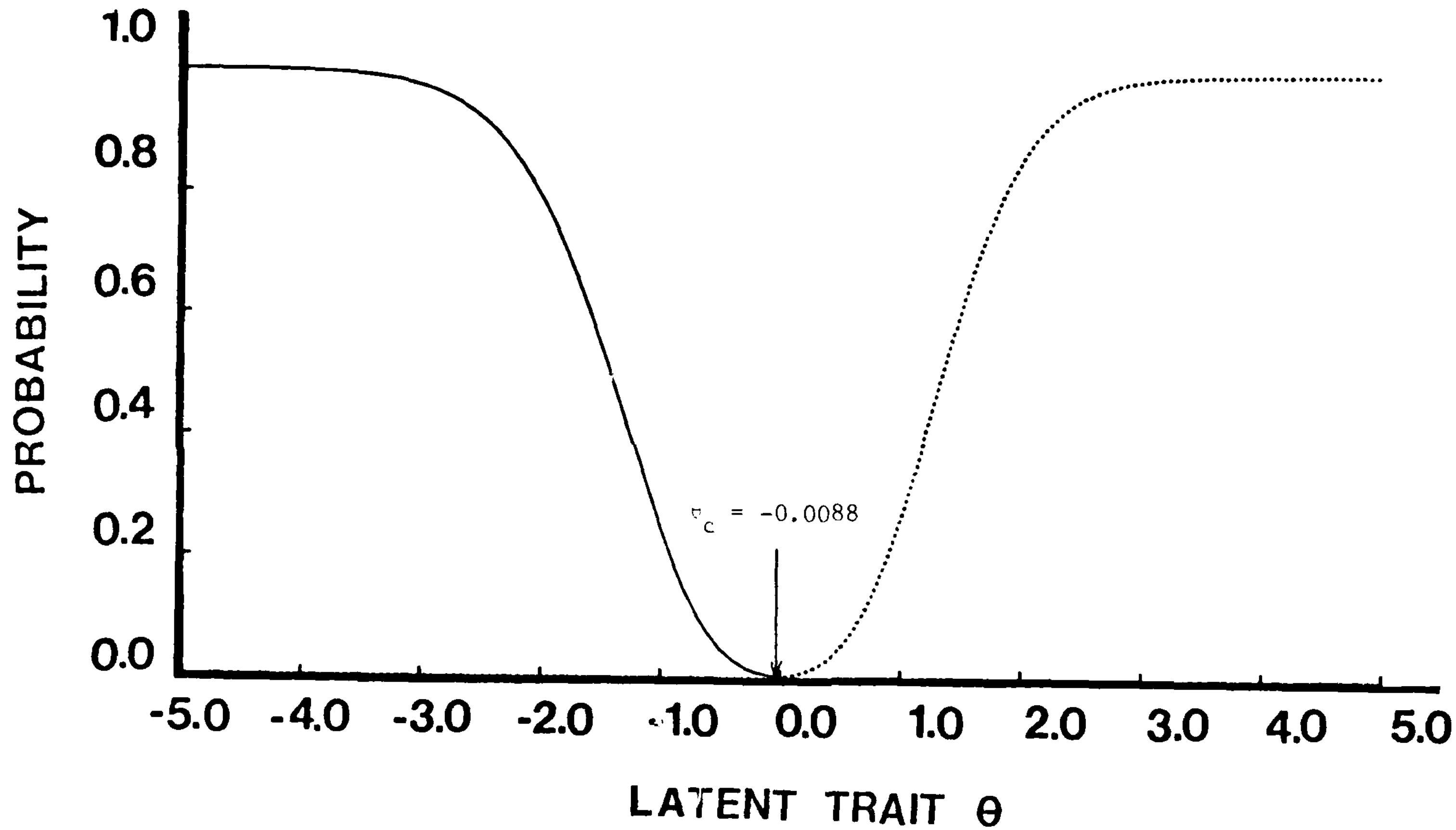


FIGURE 2-2

Operating Characteristics of the Two Extreme Response Patterns, (0,0,0,0,0,0,0) (Solid Line) and (1,1,1,1,1,1,1) (Dotted Line), of LIS-U , and the Position of the Critical Value θ_c .

TABLE 2-2

$\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ Obtained upon Each of the Eight Smallest Intervals of $(\underline{\theta},\bar{\theta})$, the Square Root of the Test Information at Each Endpoint of Each Interval, and the Upper Bound of the Discrepancies between the Regression of θ^*_V on θ and θ Itself, for LIS-U.

$\underline{\theta},\bar{\theta}$	$\sqrt{I(\underline{\theta})}$	$\sqrt{I(\bar{\theta})}$	$\theta^*_{V\text{-min}}$	$\theta^*_{V\text{-max}}$	upper bound of discrepancies
±1.50	1.439	1.494	-1.479	1.522	0.20
±1.75	1.197	1.248	-1.647	1.656	0.26
±2.00	0.982	1.001	-1.793	1.776	0.33
±2.25	0.801	0.773	-1.925	1.892	0.42
±2.50	0.648	0.574	-2.051	2.008	0.53
±3.00	0.407	0.294	-2.295	2.241	0.77
±3.50	0.233	0.147	-2.536	2.480	1.02
±4.00	0.121	0.075	-2.779	2.723	1.28

of θ^*_V and the true value of θ . We can see in this table that, even for the smallest interval, (-1.50, 1.50) , the square root of the test information assumes values as low as 1.44 and 1.49 , respectively, at the two endpoints of the interval, and yet the upper bound of the discrepancies is considerably low for the first, say, five intervals. This fact indicates that, in spite of the relatively low amounts of test information of LIS-U , the introduction of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ has succeeded in providing us with an approximately unbiased estimator, i.e., the modified maximum likelihood estimator θ^*_V , which can be used for a fairly large interval of θ . Since the least finite value of the maximum likelihood estimate is -1.3167 for the response pattern, (0,0,0,1,0,0,0) , and the greatest finite value is 1.3028 for (1,1,1,0,1,1,1) , the pair of values, -1.479 and 1.522 , obtained for $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ upon the interval, (-1.50, 1.50) , sounds reasonable enough. We could expand the interval, however, by using one of the other pairs of estimates, which are larger in absolute values than the above values of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, and, with the trade-off of the length of interval against the upper bound of the discrepancies between the regression and the true value of θ , we may conclude that we should use one of the first five intervals of $(\underline{\theta},\bar{\theta})$, the largest of which is (-2.50, 2.50) .

In the present study, we choose the interval (-2.25, 2.25) for $(\underline{\theta},\bar{\theta})$, which provides us with -1.925 and 1.892 for $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, respectively. As we can see in Table 2-2, this selection assures us that the conditional expectation of our estimate, given ability

θ , does not differ from the true value of θ by more than 0.42 , at any point of the interval of θ . Figure 2-3 presents the regression of θ^*_V on θ , which is given by

$$(2.16) \qquad E(\theta^*_V|\theta) = \sum_V \theta^*_V P_V(\theta) ,$$

by a solid curve. In the same figure, also presented is the standard error of estimate, which is defined by

$$(2.17) \qquad \{E[(\theta^*_V - E(\theta^*_V|\theta)]^2\}^{1/2} = \{\sum_V [\theta^*_V - E(\theta^*_V|\theta)]^2 P_V(\theta)\}^{1/2} ,$$

and is plotted, vertically, by dots in both negative and positive directions from the regression. There is a straight line with forty-five degrees from the abscissa of the figure, which indicates the unbiasedness, and, hereafter, we shall call it the unbiasedness line. The reciprocal of the square root of the test information function, which is usually considered as the standard error in the maximum likelihood estimation, is also plotted by dotted lines, vertically, in both negative and positive directions from the unbiasedness line. It is interesting to note that, while the standard error for the maximum likelihood estimate increases as the regression diverts from the center of the interval, the counterpart for the modified maximum likelihood estimate, θ^*_V , decreases, and the two dotted curves for the latter and the regression itself converge to $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, respectively, as θ tends to negative and positive infinities. We can also see that, for almost

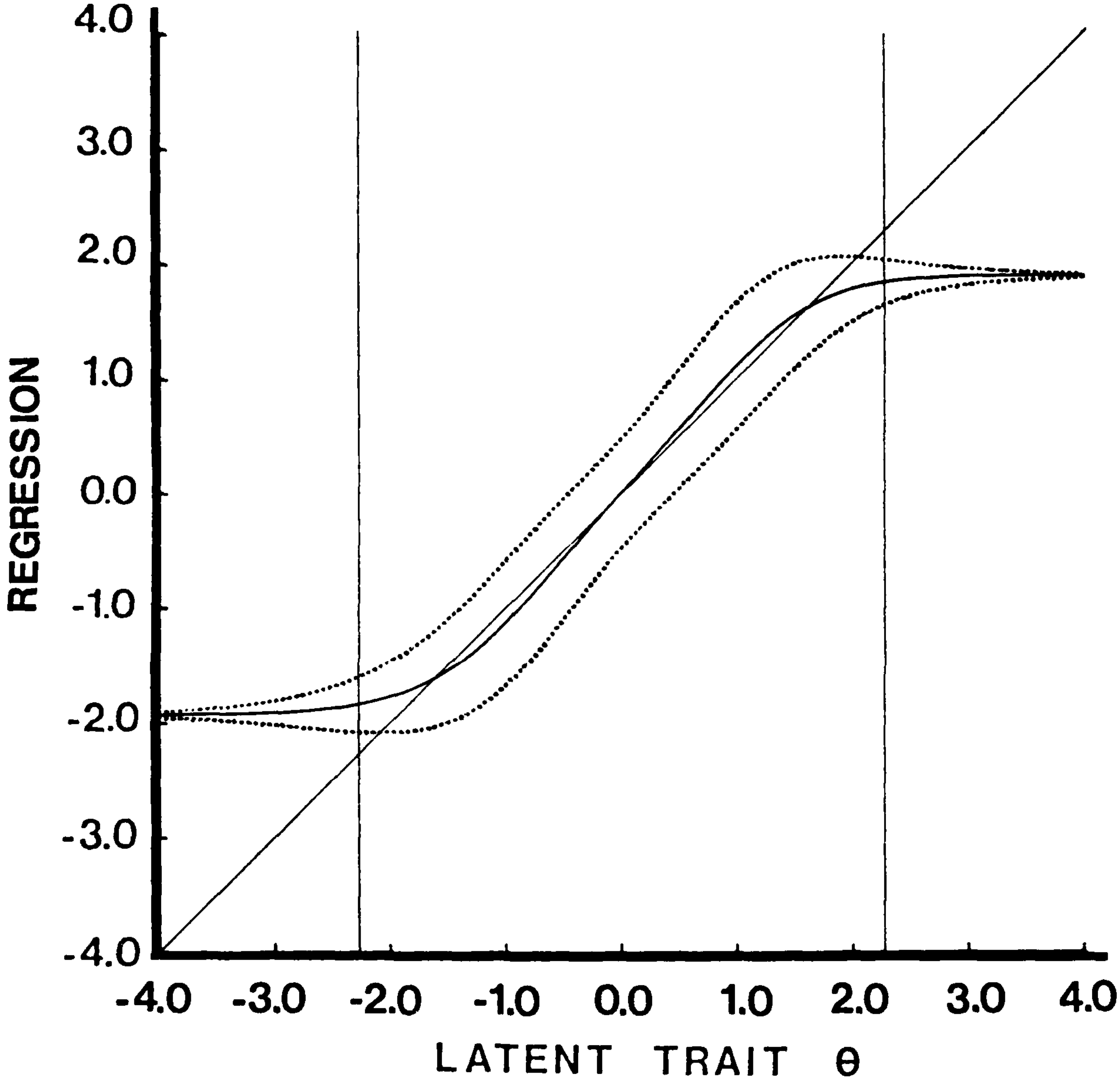


FIGURE 2-3

Regression of the Modified Maximum Likelihood Estimate θ^*_V (Solid Curve) Based upon the Interval, $-2.25 \leq \theta \leq 2.25$, on Ability θ, for LIS-U. The Standard Error of Estimate Is Plotted by Dots in Both Vertically Positive and Negative Directions from the Regression, As a Function of Ability θ.

the entire range of the interval, the unbiasedness line lies within the vertical interval of the standard error for the modified maximum likelihood estimate, θ^*_V .

It has been pointed out (Samejima, RR-80-3) that, unlike the modified maximum likelihood estimate, θ^*_V , any Bayesian estimate involves the bias caused by the prior, which contradicts the principle of the objectivity of testing. Let $\hat{\theta}_V$ be the Bayes modal estimate for a specific response pattern V . This estimate is defined as the value of θ at which the function $B_V(\theta)$, which is given by

$$(2.18) \qquad B_V(\theta) = f(\theta)\ P_V(\theta) \ ,$$

assumes the maximal value, where $f(\theta)$ is the density function of θ , or the prior. Figure 2-4 presents the regression of the Bayes modal estimate $\hat{\theta}_V$, which is obtained by replacing θ^*_V by $\hat{\theta}_V$ in (2.16), and the vertical interval similar to the one in Figure 2-3, with the standard error of estimate obtained by replacing θ^*_V by $\hat{\theta}_V$ in (2.17), by solid and dotted lines, respectively. Comparison of this figure with Figure 2-3 reveals that the vertical interval for the Bayes modal estimate $\hat{\theta}_V$ contains the unbiasedness line only for a much smaller interval of θ , i.e., approximately $(-1.38, 1.41)$, and outside this interval both the vertical interval and the regression converge quickly to $\hat{\theta}_{V\text{-min}} = -1.3617$ and $\hat{\theta}_{V\text{-max}} = 1.3542$, respectively, as θ tends to negative and positive infinities.

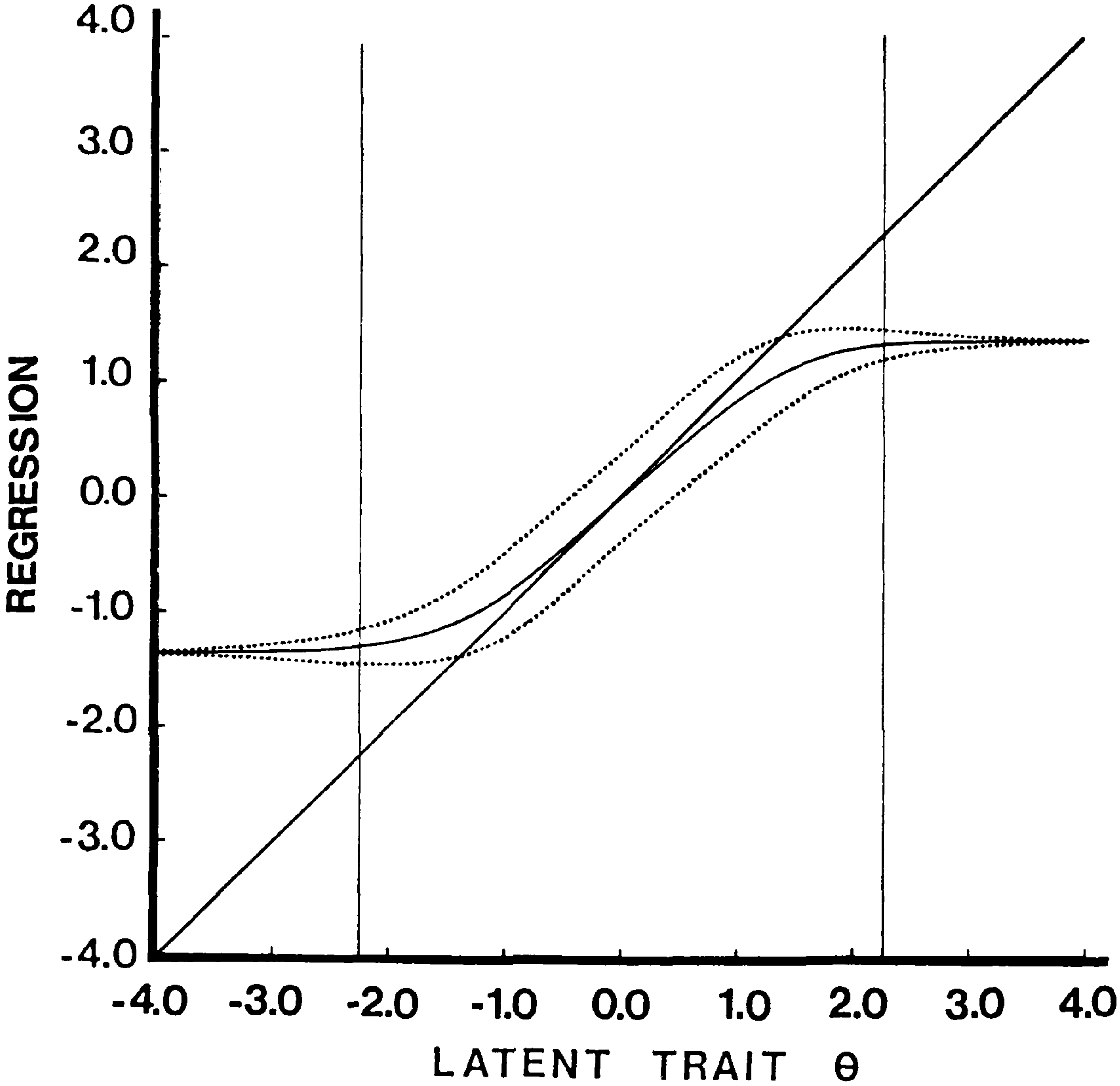


FIGURE 2-4

Regression of the Bayes Modal Estimate $\hat{\theta}_V$ with the Prior n(0,1) (Solid Curve) on Ability θ, for LIS-U. The Standard Error of Estimate Is Plotted by Dots in Both Vertically Positive and Negative Directions from the Regression, As a Function of Ability θ.

We shall observe another estimate and its regression on ability θ and the similar vertical interval for the standard error of estimate. This is Bayes estimate, μ'_{1V}, with the same prior, $n(0,1)$, as we used for the Bayes modal estimate, $\hat{\hat{\theta}}_V$. This estimator is defined by

$$(2.19) \qquad \mu'_{1V} = \int_{-\infty}^{\infty} \theta \, f_V(\theta) \, d\theta \, ,$$

where $f_V(\theta)$ is the density function of θ for the subgroup of examinees whose response patterns are uniformly V, which is given by

$$(2.20) \qquad f_V(\theta) = f(\theta) \, P_V(\theta) \, [\int_{-\infty}^{\infty} f(\theta) \, P_V(\theta) \, d\theta]^{-1} \, .$$

Figure 2-5 presents, for the Bayes estimate, μ'_{1V}, a set of functions similar to those which we have observed for both the modified maximum likelihood estimate, θ^*_V, and the Bayes modal estimate, $\hat{\hat{\theta}}_V$, in Figures 2-3 and 2-4, respectively. They are the regression obtained by replacing θ^*_V by μ'_{1V} in (2.16) and the interval based upon the standard error of estimate obtained by the similar replacement in (2.17). We can see that this set of results is very much like the one we obtained for the Bayes modal estimate with the same prior, $n(0,1)$, which we have observed in Figure 2-4. The interval of θ for which the vertical interval of the standard error of estimate includes the unbiasedness line is approximately $(-1.40, 1.28)$, and outside of this small interval of θ the three curves converge quickly to $\mu'_{1V\text{-}min} = -1.3764$ and $\mu'_{1V\text{-}max} = 1.2695$, respectively.

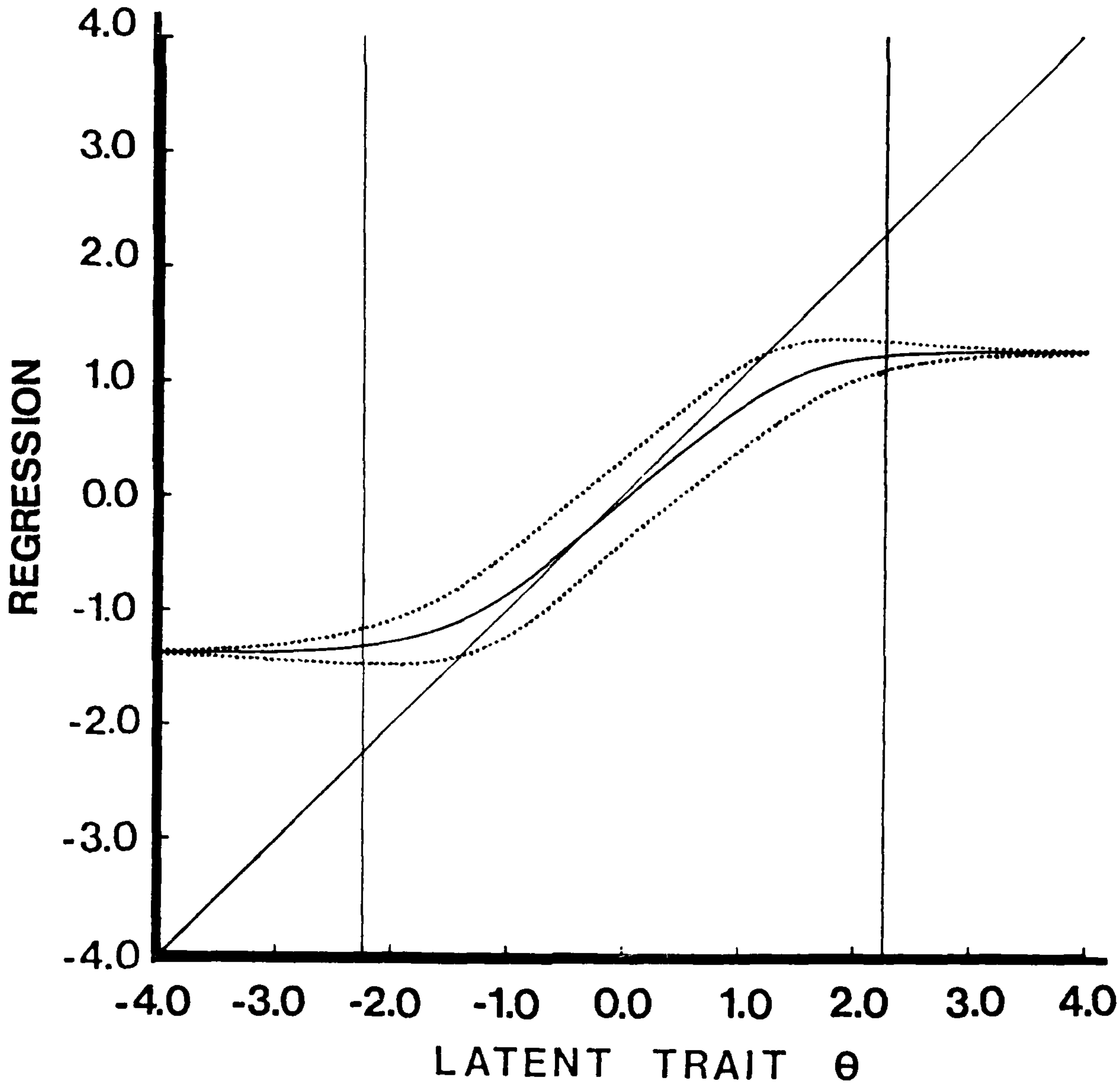


FIGURE 2-5

Regression of the Bayes Modal Estimate θ'_{1V} with the Prior n(0,1) (Solid Curve) on Ability θ, for LIS-U. The Standard Error of Estimate Is Plotted By Dots in Both Vertically Positive and Negative Directions From the Regression, As a Function of Ability θ.

The similarity observed in the above two results makes us wonder if we can define an estimator which is population-free, and has an analogous meaning to the maximum likelihood estimator as Bayes estimator has to the Bayes modal estimator. Such an estimator may be a better estimator than the modified maximum likelihood estimator, θ^*_V , or may not. To find it out, we shall introduce an estimator defined by (2.19) and (2.20), where $f(\theta)$ is a uniform density function for the interval, $(\underline{\theta},\bar{\theta})$. Let $\mu^{*\prime}_{1V}$ denote this Bayes estimate with the uniform prior. We can write from (2.19) and (2.20) that

$$(2.21) \qquad \mu^{*\prime}_{1V} = \int_{\underline{\theta}}^{\bar{\theta}} \theta\, P_V(\theta)\, d\theta \left[\int_{\underline{\theta}}^{\bar{\theta}} P_V(\theta)\, d\theta \right]^{-1} .$$

Note that (2.21) includes only the conditional probability of response pattern V , given ability θ , and, therefore, is population-free.

Figure 2-6 presents the regression of $\mu^{*\prime}_{1V}$ on θ , which is obtained by replacing θ^*_V by $\mu^{*\prime}_{1V}$ in (2.16), and the vertical interval of the standard error of estimate, which is obtained by replacing θ^*_V by $\mu^{*\prime}_{1V}$ in (2.17), by solid and dotted lines, respectively. As was expected, this set of results is similar to the one obtained upon the modified maximum likelihood estimate, θ^*_V , which is shown as Figure 2-3. A close observation reveals, however, that the interval of θ for which the vertical interval of the standard error of estimate includes the unbiasedness line is somewhat smaller, i.e., (-1.74, 1.76) , compared with (-2.08, 2.06) for the modified maximum likelihood estimate, although

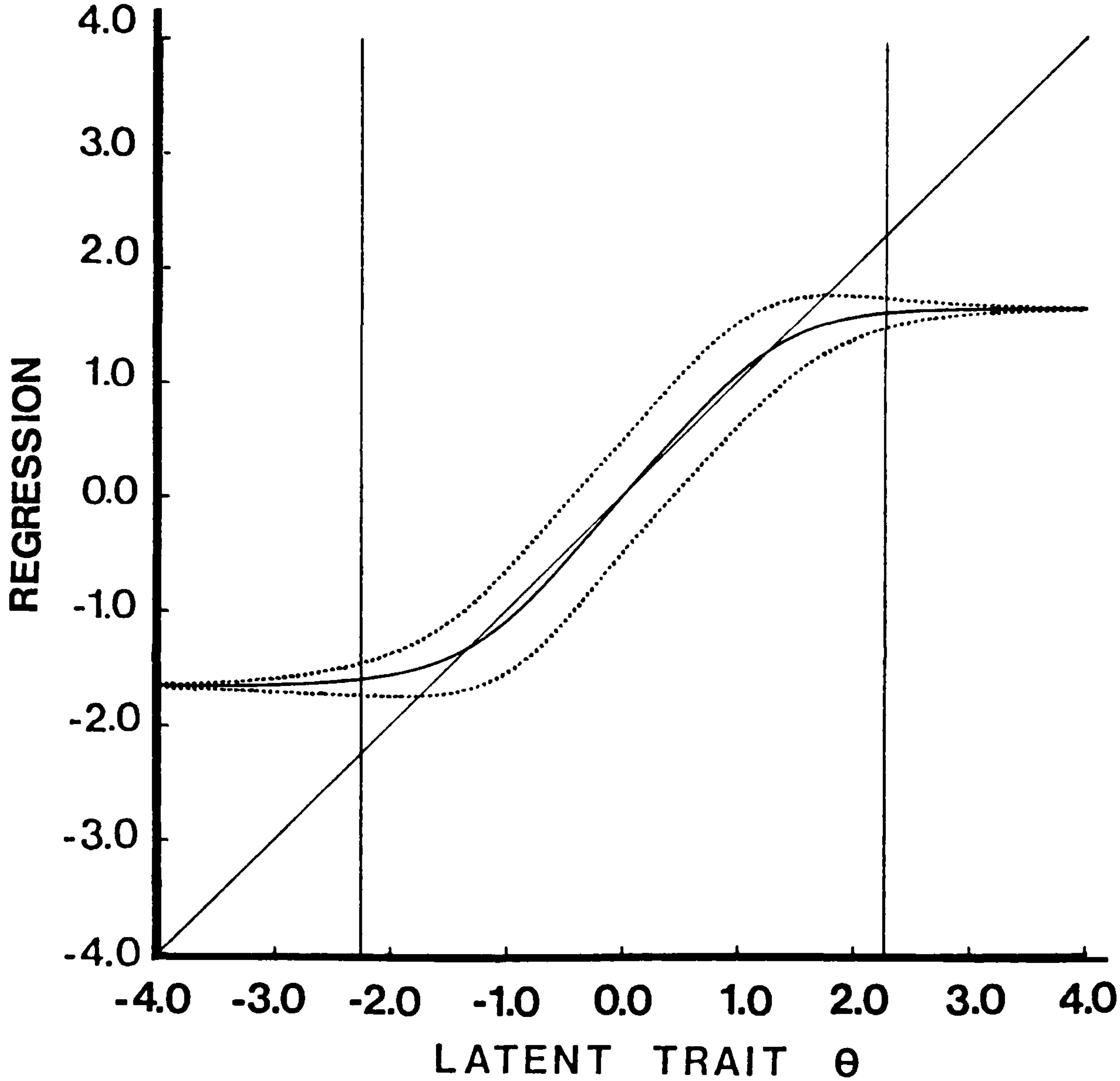


FIGURE 2-6

Regression of the Bayes Estimate $\hat{\mu}^{*\prime}_{IV}$ with the Uniform Prior for $-2.25 \le \theta \le 2.25$, on Ability θ, for LIS-U. The Standard Error of Estimate Is Plotted By Dots in Both Vertically Positive and Negative Directions from the Regression, As a Function of Ability θ.

this interval is substantially larger than the intervals found for the two Bayesian estimates. It is also noted by comparing Figures 2-3 and 2-6 that, for the entire interval of $(\underline{\theta},\bar{\theta})$, the regression of the modified maximum likelihood estimate, θ^*_V , tends to be closer to the unbiasedness line than the regression of $\mu^{*\prime}_{1V}$. The values of $\mu^{*\prime}_{1V}$ for the two extreme response patterns, V-min and V-max , turned out to be more conservative than those of $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, i.e., -1.6515 vs. -1.9254 and 1.6430 vs. 1.8923 , respectively.

The four sets of estimates for the total one hundred and twenty-eight response patterns of LIS-U are shown in Appendix, as Table A-1. They are also pairwisely plotted in six scatter diagrams, and presented as Figures 2-7 through 2-12. As is expected, the scatter diagram of $\mu^{*\prime}_{1V}$ plotted against the modified maximum likelihood estimate, θ^*_V , and that of the Bayes modal estimate, $\hat{\hat{\theta}}_V$, plotted against the Bayes estimate, $\mu^{\prime}_{1V}$, which are shown as Figures 2-7 and 2-12, respectively, are almost on the line with forty-five degrees from the abscissa, while in the other four combinations of estimates they are consistently and substantially deviated from this line. It is interesting to note that, in Figure 2-7, for all the other one hundred and twenty-six response patterns excluding V-min and V-max , the values of θ^*_V and $\mu^{*\prime}_{1V}$ are very close to each other, while for these two extreme response patterns the latter are substantially smaller in absolute values than the former.

From these results, we can say that the modified maximum likelihood

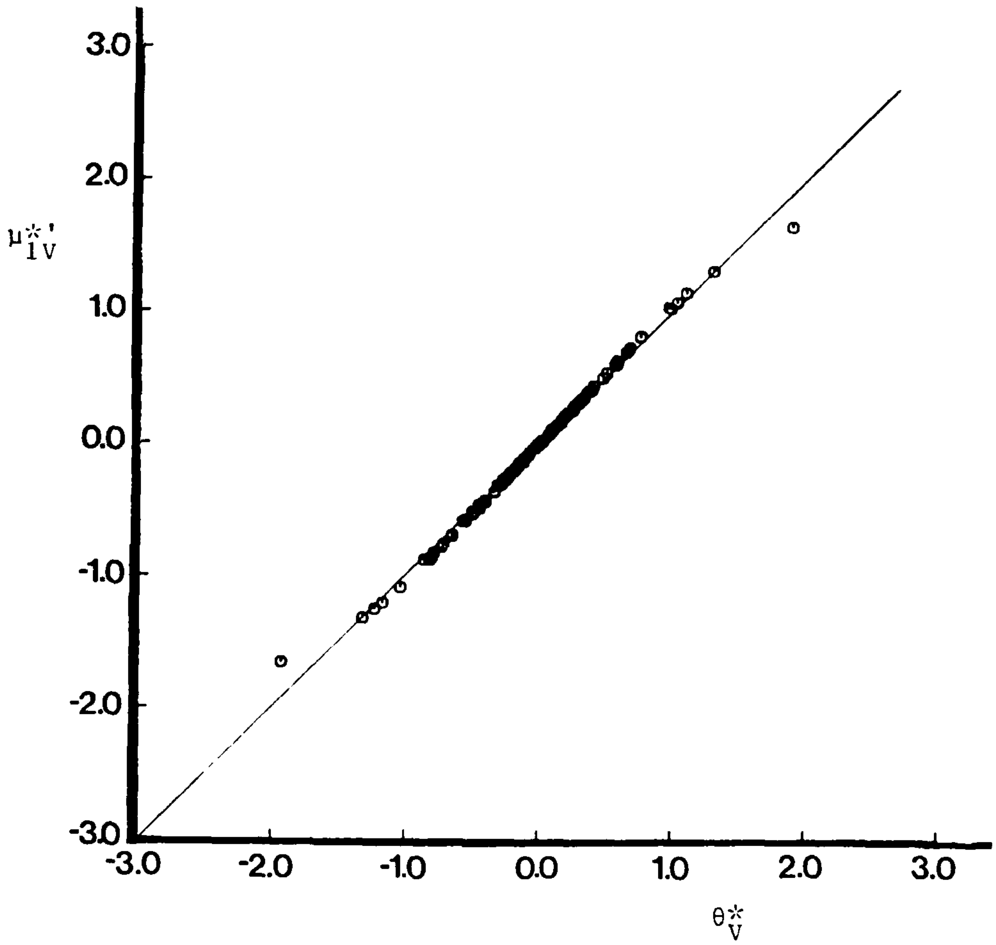


FIGURE 2-7

Bayes Estimate with the Uniform Prior, $\mu^{*'}_{IV}$, Plotted against the Modified Maximum Likelihood Estimate, θ^*_V, for the One Hundred and Twenty-eight Possible Response Patterns of LIS-U.

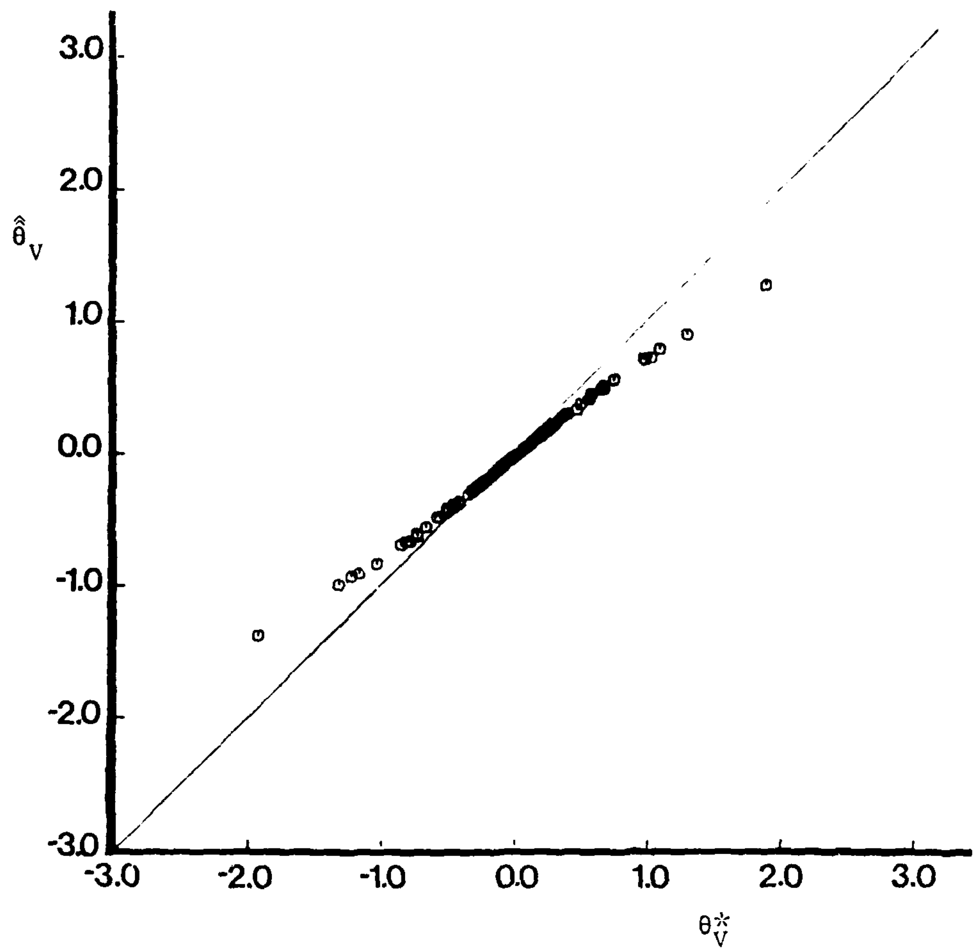


FIGURE 2-8

Bayes Modal Estimate, $\hat{\hat{\theta}}_V$, with the Prior $n(0,1)$, Plotted against the Modified Maximum Likelihood Estimate, θ^*_V , for the One Hundred and Twenty-eight Possible Response Patterns of LIS-U.

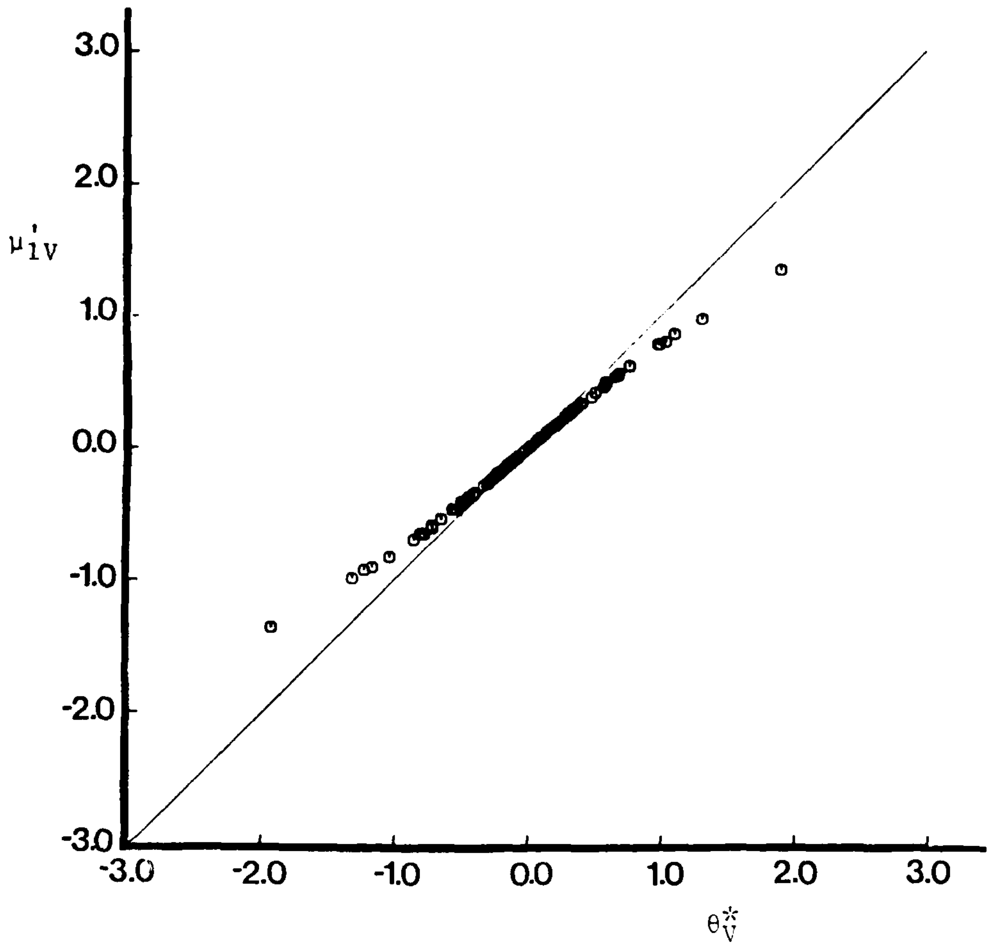


FIGURE 2-9

Bayes Estimate , μ'_{1V} , with the Prior n(0,1) , Plotted against the Modified Maximum Likelihood Estimate, θ^*_V , for the One Hundred and Twenty-eight Possible Response Patterns of LIS-U.

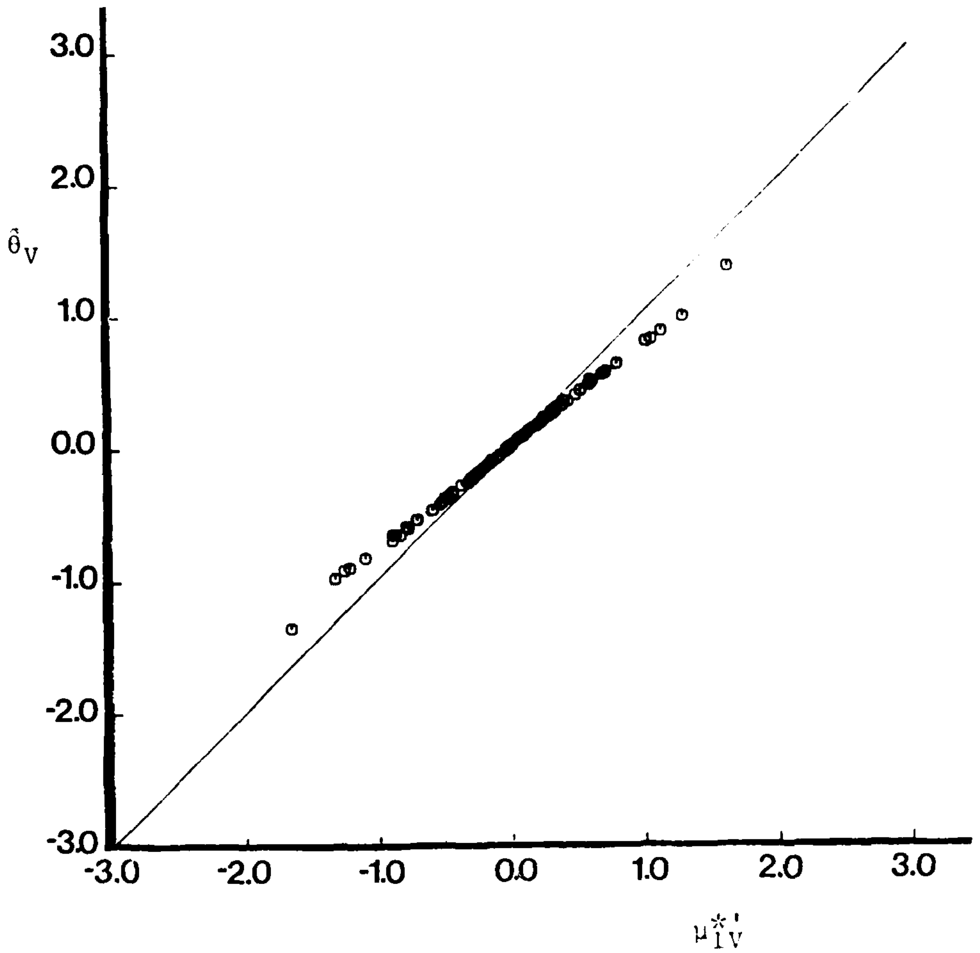


FIGURE 2-10

Bayes Modal Estimate, $\hat{\hat{\theta}}_V$, with the Prior n(0,1) , Plotted against the Bayes Estimate with the Uniform Prior, $\mu^{*\prime}_{IV}$, for the One Hundred and Twenty-eight Possible Response Patterns of LIS-U.

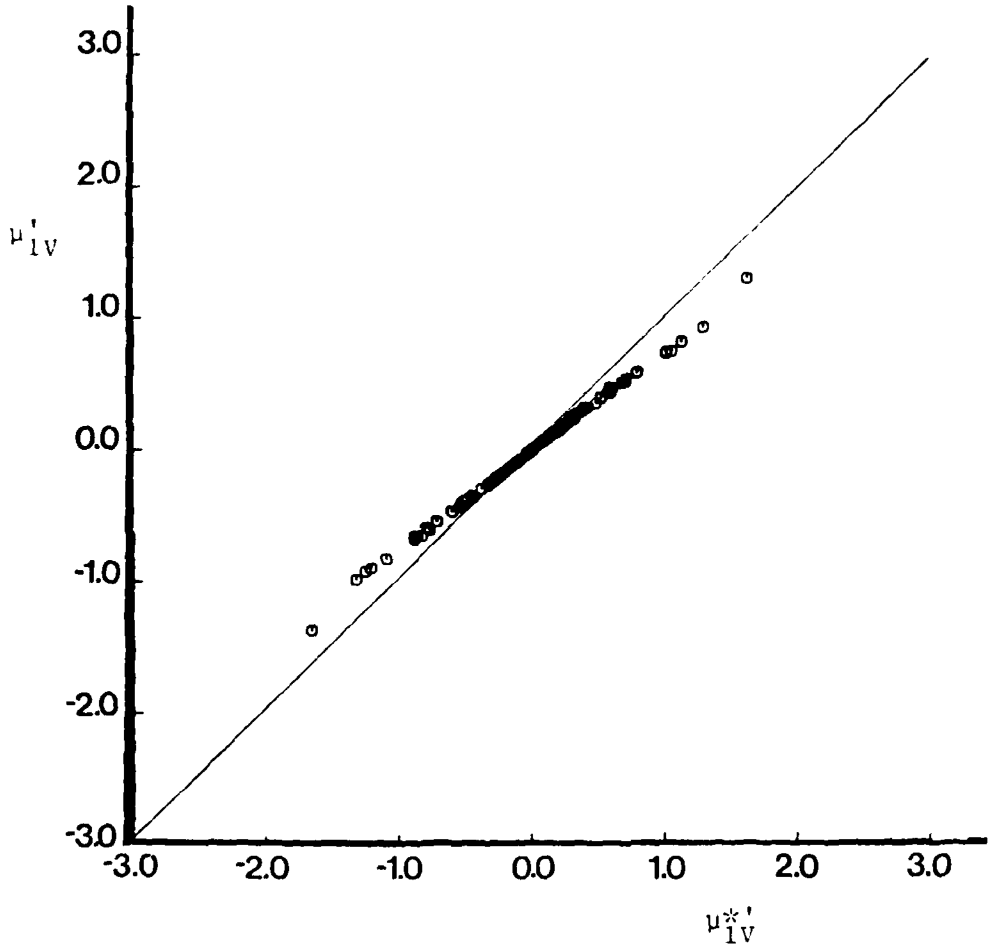


FIGURE 2-11

Bayes Estimate, μ'_{1V} , with the Prior $n(0,1)$, Plotted against the Bayes Estimate with the Uniform Prior, $\mu^{*\prime}_{1V}$, for the One Hundred and Twenty-eight Possible Response Patterns of LIS-U.

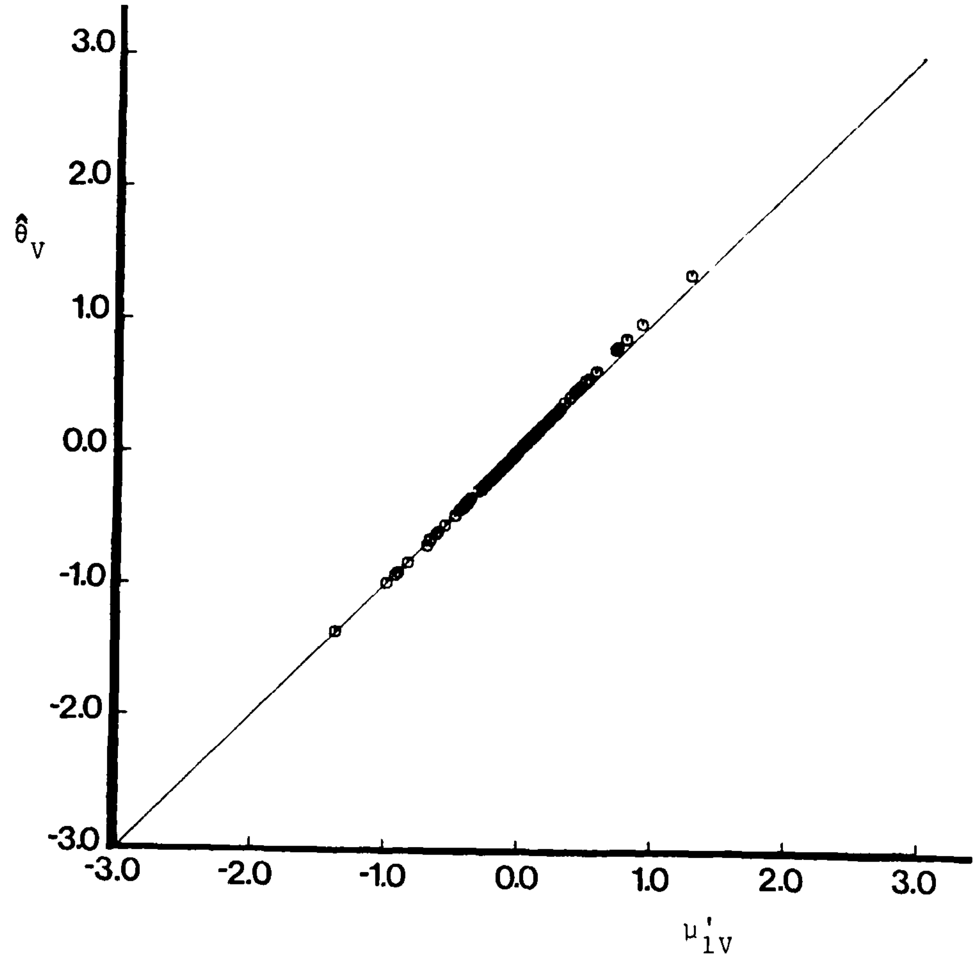


FIGURE 2-12

Bayes Modal Estimate, $\hat{\theta}_V$, with the Prior n(0,1) , Plotted against the Bayes Estimate, μ'_{1V} , with the Prior n(0,1) , for the One Hundred and Twenty-eight Possible Response Patterns of LIS-V.

estimator, θ^*_V , provides us with a better approximation to the unbiasedness and is a better estimator than $\mu^{*\prime}_{1V}$, although the latter is also population-free and is a much better estimator than the Bayesian types of estimators in satisfying the principle of objective testing.

III Sample Statistic Versions of the Alternative Estimators for the Two Extreme Response Patterns

The introduction of the two alternative estimators for V-min and V-max and the resultant modified maximum likelihood estimate, θ^*_V , has enhanced the usefulness of relatively short and less informative tests, without sacrificing the objectivity of testing. When the number of items is as small as seven and all items are binary items, as is the case with LIS-U , the computation of θ^*_{V-min} and θ^*_{V-max} is relatively easy, owing to the fact that the number of all possible response patterns is as small as 128 . Note, however, that the increase in the number of items, and/or in the number of item scores for each item, will soon make it practically impossible to compute these two substitute estimates, since the number of all possible response patterns will increase by gigantic steps. For example, if a test has ten binary items instead of seven, the number of all possible response patterns will be 1,024 ; if a test has seven three-item-score-category items, the number of all possible response patterns will be 2,187 ; if a test has fifteen three-item-score-category items, it will be as large as 14,348,907 !

For the reason described in the preceding paragraph, it is necessary that we should invent some device in dealing with the situation in which the number of all possible response patterns is too large for us to compute θ^*_{V-min} and θ^*_{V-max} directly. By virtue of the availability of electronic computers and the Monte Carlo method, this can be done by introducing the sample statistic

versions of the two estimators.

Let N be the number of examinees who were selected randomly from the uniform distribution for the interval of θ , $(\underline{\theta},\bar{\theta})$. Let N_L denote the number of examinees who belong to the above sample and whose levels of ability are lower than the critical value θ_c , and N_H be of that of those whose ability levels are higher than, or equal to, θ_c . Thus we can write

$$(3.1) \qquad N = N_L + N_H \ .$$

Let N_{LV} and N_{HV} denote the numbers of examinees who obtained the response pattern V , in the above two subgroups of the sample, respectively. Thus we have

$$(3.2) \qquad \begin{cases} N_L = \sum_V N_{LV} \\ N_H = \sum_V N_{HV} \ . \end{cases}$$

It can be seen that the sample statistic corresponding to $\int_{\underline{\theta}}^{\theta_c} P_V(\theta)\, d\theta$ in the formula (2.14) is $N_{LV}(\theta_c - \underline{\theta})\, N_L^{-1}$, and also the one for $\int_{\theta_c}^{\bar{\theta}} P_V(\theta)\, d\theta$ is $N_{HV}\ (\bar{\theta} - \theta_c)\ N_H^{-1}$. Substituting these sample statistics into (2.14) and rearranging, we obtain $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$ such that

$$(3.3) \qquad \begin{cases} \hat{\theta}^*_{V\text{-min}} = [\frac{1}{2}(\theta_c + \underline{\theta})\ N_L - \sum\limits_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V\ N_{LV}]\ N_{LV\text{-min}}^{-1} \\ \hat{\theta}^*_{V\text{-max}} = [\frac{1}{2}(\bar{\theta} + \theta_c)\ N_H - \sum\limits_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\theta}_V\ N_{HV}]\ N_{HV\text{-max}}^{-1} \ , \end{cases}$$

where $N_{LV\text{-}min}$ and $N_{HV\text{-}max}$ are the numbers of examinees who belong to the lower subgroup and obtained the response pattern V-min , and those who belong to the upper subgroup and obtained the response pattern V-max , respectively.

It can be seen that $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$, which were defined in the preceding paragraph, are consistent, or converge in probability to $\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$, respectively, as the sample sizes increase. In other words, if N_L , N_H , $N_{LV\text{-}min}$ and $N_{HV\text{-}max}$ are large enough, the probabilities with which $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$ assume values within the vicinities of $\theta^*_{V\text{-}min}$ and $\theta^*_{V\text{-}max}$, respectively, will be very high. Although the two numbers, $N_{LV\text{-}min}$ and $N_{HV\text{-}max}$, also depend upon the choice of the interval, $(\underline{\theta},\bar{\theta})$, by virtue of the Monte Carlo method, we can control the two sample sizes, N_L and N_H , as we wish.

A procedure with which we may obtain $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$, which are defined by (3.3), can be summarized as follows.

(1) Determine the interval, $(\underline{\theta},\bar{\theta})$.

(2) Obtain the critical value, θ_c .

(3) Determine the sample size, N , which makes both N_L and N_H large enough for our purpose.

(4) Produce the ability levels of these N hypothetical subjects from the uniform distribution for the interval, $(\underline{\theta},\bar{\theta})$. This can be done either by the Monte Carlo method, or by placing the N examinees at equally spaced points in the entire

interval, $(\underline{\theta}, \bar{\theta})$, or using one of its variations.

(5) Calibrate by the Monte Carlo method a response pattern for each of the N hypothetical examinees with respect to the n test items of our test.

(6) Find out the two frequencies, N_{LV} and N_{HV} , for each response pattern V .

(7) Obtain the maximum likelihood estimate $\hat{\theta}_V$ for each response pattern whose frequencies, N_{LV} and N_{HV} , are not both zero, excluding V-min and V-max .

(8) Use the above results in (3.3), and compute $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$.

Note that the probabilities with which we obtain positive frequencies for $N_{LV\text{-max}}$ and $N_{HV\text{-min}}$ are both negligibly small, and this fact can be used as a checking process.

For the purpose of illustration, we shall use the five hundred hypothetical examinees, who have been used repeatedly in our previous studies of estimating the operating characteristics of the discrete item responses (Samejima, 1977c, RR-77-1, RR-78-1, RR-78-2, RR-78-3, RR-78-4, RR-78-5, RR-78-6, RR-80-2, RR-80-4), and Subtest 3, which is one of the subtests from the Old Test of thirty-five test items with three item score categories for each item. These hypothetical examinees are placed at the one hundred points of ability θ , which start from -2.475 and equally spaced by the distance, 0.05 , with

five examinees positioned together at each point. For this reason, they can be considered as a representative sample from the uniformly distributed population for the interval of θ , (-2.5, 2.5) . These fifteen test items of Subtest 3 follow the normal ogive model, whose operating characteristic is given by (2.2) . The item discrimination parameter, a_g , and the two item response difficulty parameters, b_{x_g} for $x_g=1$ and $x_g=2$, of these items of Subtest 3 are presented in Table 3-1. The square root of the test information function of Subtest 3, which is computed through (2.12), is drawn by a solid line in Figure 3-1. We can see that the function is uni-modal, which means that Subtest 3 is more informative around the middle of the interval of θ , (-2.5, 2.5), and less informative as θ diverts from the middle. Figure 3-2 presents the two operating characteristics, $P_{V\text{-}min}(\theta)$ and $P_{V\text{-}max}(\theta)$, by solid and dotted lines, respectively, together with the critical value, θ_c , which equals -0.4146 . Unlike the one we used for LIS-U in the preceding chapter, this value does <u>not</u> make the product of the two operating characteristics maximal, but is more deviated toward the negative side. And yet both $P_{V\text{-}min}(\theta)$ and $P_{V\text{-}max}(\theta)$ are practically zero at this point of θ .

Since the response pattern for each of the five hundred hypothetical examinees with respect to the Old Test was already calibrated (Samejima, 1977c), we have used the subset of this response pattern for the fifteen items of Subtest 3. The maximum likelihood estimate for each response pattern was obtained by using the basic functions, $A_{x_g}(\theta)$, and (2.6). It turned out that only fourteen examinees out of five hundred obtained

TABLE 3-1

Item Discrimination Parameter, a_g, and Item Response Difficulty Parameters, b_{x_g}, for $x_g = 1$ and $x_g = 2$, for Each of the Thirty-Five Test Items of the Old Test. The Fifteen Test Items Which Belong to Subtest 3 Are Marked with Crosses.

Item g	a_g	b_1	b_2	Subtest 3
1	1.8	-4.75	-3.75	
2	1.9	-4.50	-3.50	
3	2.0	-4.25	-3.25	
4	1.5	-4.00	-3.00	
5	1.6	-3.75	-2.75	
6	1.4	-3.50	-2.50	
7	1.9	-3.00	-2.00	
8	1.8	-3.00	-2.00	
9	1.6	-2.75	-1.75	
10	2.0	-2.50	-1.50	
11	1.5	-2.25	-1.25	x
12	1.7	-2.00	-1.00	x
13	1.5	-1.75	-0.75	x
14	1.4	-1.50	-0.50	x
15	2.0	-1.25	-0.25	x
16	1.6	-1.00	0.00	x
17	1.8	-0.75	0.25	x
18	1.7	-0.50	0.50	x
19	1.9	-0.25	0.75	x
20	1.7	0.00	1.00	x
21	1.5	0.25	1.25	x
22	1.8	0.50	1.50	x
23	1.4	0.75	1.75	x
24	1.9	1.00	2.00	x
25	2.0	1.25	2.25	x
26	1.6	1.50	2.50	
27	1.7	1.75	2.75	
28	1.4	2.00	3.00	
29	1.9	2.25	3.25	
30	1.6	2.50	3.50	
31	1.5	2.75	3.75	
32	1.7	3.00	4.00	
33	1.8	3.25	4.25	
34	2.0	3.50	4.50	
35	1.4	3.75	4.75	

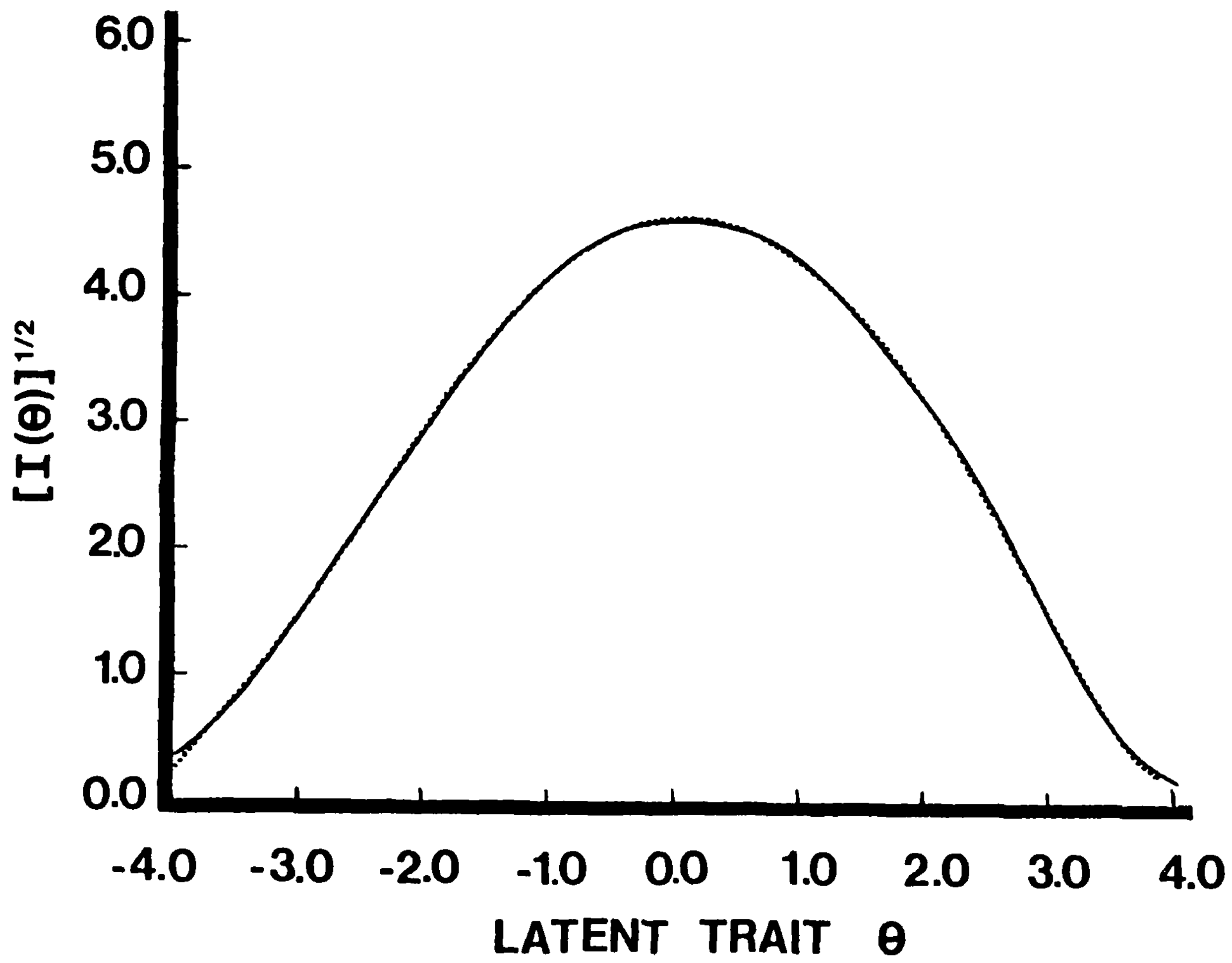


FIGURE 3-1

Square Root of Test Information Function (Solid Line), and Its Approximation by the Polynomial of Degree 7 (Dotted Line), of Subtest 3.

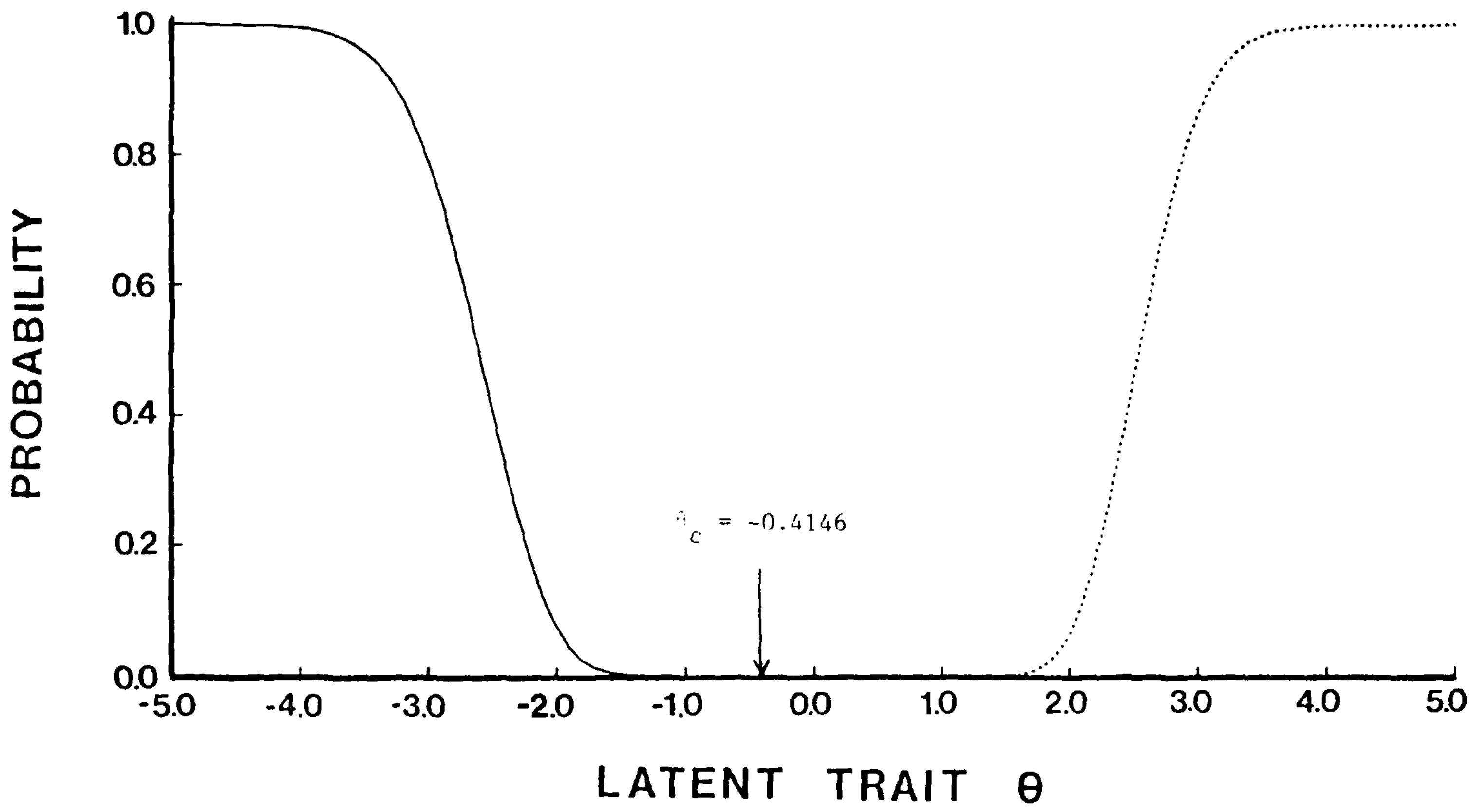


FIGURE 3-2

Operating Characteristics of the Two Extreme Response Patterns, $P_{V-min}(\theta)$ (Solid Line) and $P_{V-max}(\theta)$ (Dotted Line), of Subtest 3, and the Position of the Critical Value, θ_c .

the response pattern, V-min , and twelve obtained the response pattern, V-max . These relatively small numbers of examinees who obtained either one of the two extreme response patterns indicate that, for Subtest 3, the interval, (-2.5, 2.5) , is still too conservative to use as $(\underline{\theta},\bar{\theta})$, and it may be expanded further. The two sample sizes, N_L and N_H , are 210 and 290 , respectively. Substituting these values, together with $N_{LV\text{-}min} = 14$, $N_{HV\text{-}max} = 12$, $\underline{\theta} = -2.5$, $\bar{\theta} = 2.5$, and $\theta_c = -0.4146$, into (3.3), we obtained $\hat{\theta}^*_{V\text{-}min} = -2.31453$ and $\hat{\theta}^*_{V\text{-}max} = 2.04027$.

Let $\hat{\theta}^*_V$ denote a new estimator, which is defined by

$$
(3.4) \qquad \hat{\theta}^*_V \begin{cases} = \hat{\theta}^*_{V\text{-}min} & \text{for } V = V\text{-min} \\ = \hat{\theta}^*_{V\text{-}max} & \text{for } V = V\text{-max} \\ = \hat{\theta}_V & \text{otherwise,} \end{cases}
$$

as distinct from θ^*_V , which is defined by (2.15). Figure 3-3 presents the scatter diagram of the five hundred hypothetical examinees with respect to their ability levels, θ , and the estimate, $\hat{\theta}^*_V$. We notice that, for fixed values of θ , the values of $\hat{\theta}^*_V$ scatter more widely as θ departs from the middle of the interval, (-2.5, 2.5) , but then start having truncated distributions as θ approaches either $\underline{\theta}$ or $\bar{\theta}$. For the purpose of comparison, Figure 3-4 presents the corresponding scatter diagram of the same five hundred hypothetical examinees, with the maximum likelihood estimate $\hat{\theta}_V$, which was based upon the original Old Test, on the ordinate. In this figure, the values of the maximum likelihood estimate for fixed ability levels scatter

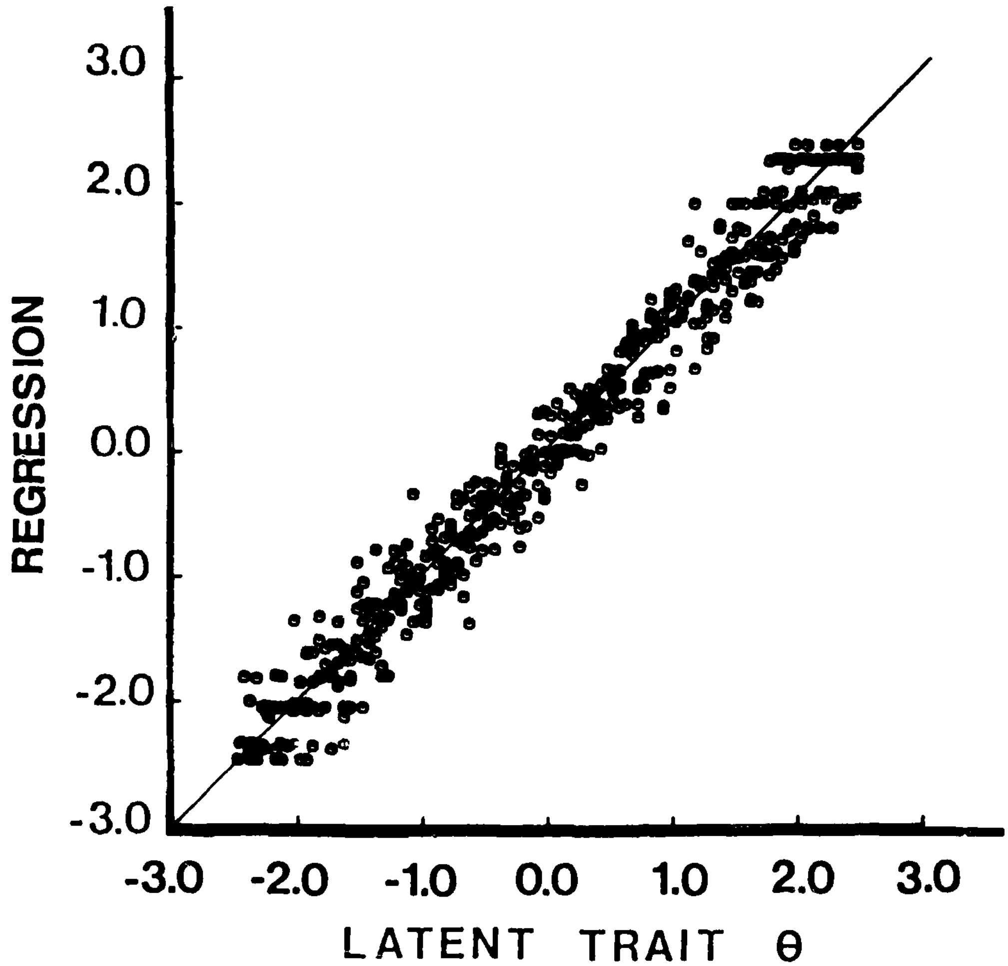


FIGURE 3-3

Scatter Diagram of $\hat{\theta}^*_V$ and θ for the Five Hundred Hypothetical Examinees, Based on Subtest 3.

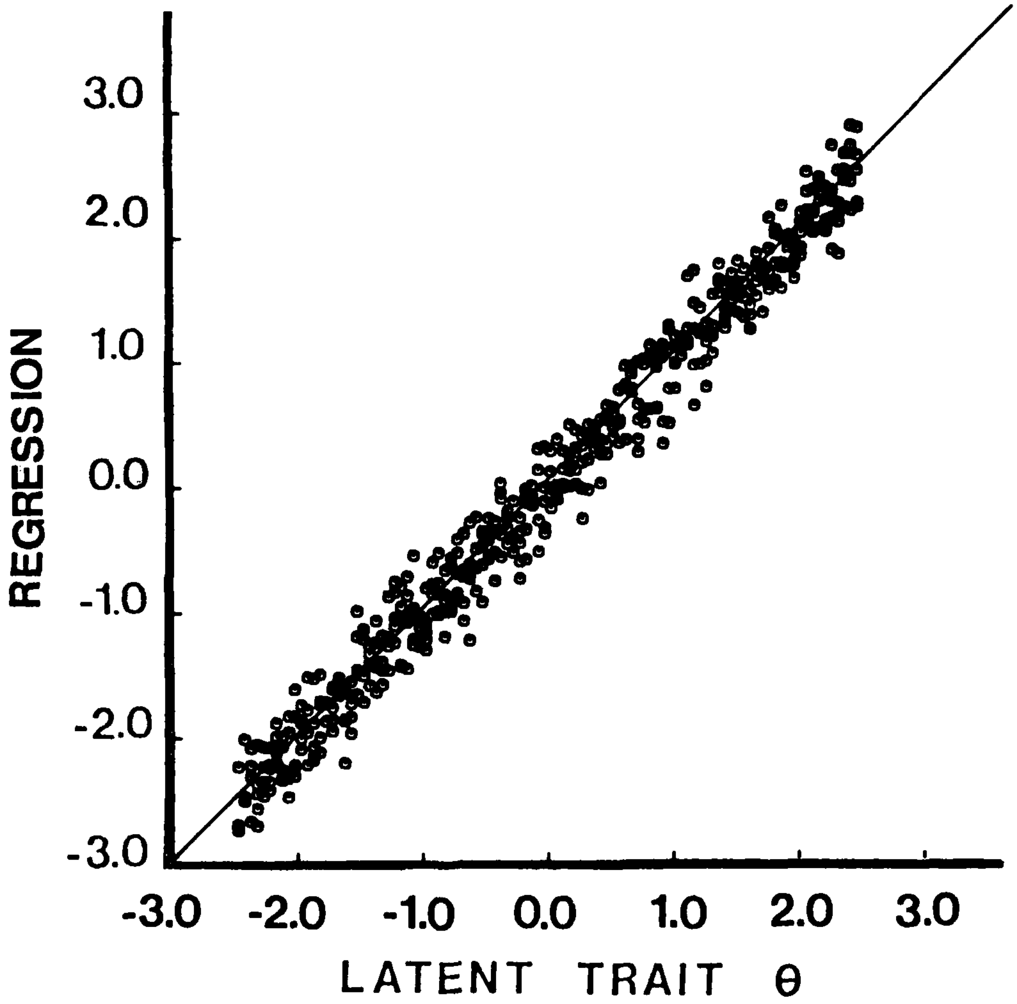


FIGURE 3-4

Scatter Diagram of the Maximum Likelihood Estimate $\hat{\theta}_V$ and Ability θ for the Five Hundred Hypothetical Examinees, Based Upon the Original Old Test.

approximately within the same range for the entire interval of θ , reflecting the fact that the test information function of the Old Test assumes approximately the same value throughout the interval of θ in question.

The sample linear regression of $\hat{\theta}_V^*$ on ability θ , or the best fitted linear function of θ which makes the sum total of the squared discrepancies of $\hat{\theta}_V^*$ minimal, turned out to be

$$(3.5) \qquad 0.99546\theta - 0.00730 \ .$$

This function of θ is shown as the straight line in Figure 3-3, which is practically indistinguishable from the unbiasedness line, or the line with forty-five degrees from the abscissa which passes the origin of the two axes, (0,0) . The corresponding sample linear regression for the scatter diagram, which is based upon the original Old Test and shown in Figure 3-4, proved to be $1.004\theta - 0.006$ (Samejima, 1977c). We can say that these two results are practically the same. The sample regression of $\hat{\theta}_V^*$ on ability θ was obtained for the one hundred ability levels, and is shown in Appendix as Figure A-1.

The mean and variance of θ for the five hundred hypothetical examinees are 0.0000 and 2.0831 , and those of $\hat{\theta}_V^*$ turned out to be -0.0073 and 2.1290 , respectively. The product-moment correlation coefficient between θ and $\hat{\theta}_V^*$ for the five hundred hypothetical examinees is found to be 0.985 .

Since we have for the uniform distribution

$$(3.6) \qquad E(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} \theta \; (\bar{\theta} - \underline{\theta})^{-1} \, d\theta = (\underline{\theta} + \bar{\theta})/2$$

and

$$(3.7) \qquad Var.(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} [\theta - E(\theta)]^2 \; (\bar{\theta} - \underline{\theta})^{-1} \, d\theta$$

$$= (\bar{\theta} - \underline{\theta})^2/12 \quad ,$$

in the present case of $\underline{\theta} = -2.5$ and $\bar{\theta} = 2.5$, we obtain

$$(3.8) \qquad E(\theta) = 0.0000$$

and

$$(3.9) \qquad Var.(\theta) = 2.0833 \quad .$$

As is expected from the way we produced the ability levels of our five hundred hypothetical examinees, the above sample mean and variance are practically the same as the population mean and variance.

When an estimator, λ , of ability θ is conditionally unbiased, or we can write

$$(3.10) \qquad E(\lambda|\theta) = \theta \; ,$$

the relationship such that

$$(3.11) \qquad E(\lambda) = E(\theta)$$

holds in general, regardless of the distribution of θ . In such a

case, the variance of λ is found out to be

$$\text{(3.12)} \qquad \mathrm{Var.}(\lambda) = \mathrm{Var.}(\theta) + E[\mathrm{Var.}(\lambda|\theta)] \geq \mathrm{Var.}(\theta)$$

(cf. Samejima, 1977c), and the product-moment correlation coefficient between θ and λ is given by

$$\text{(3.13)} \qquad \mathrm{Corr.}(\theta,\lambda) = [1-E\{\mathrm{Var.}(\lambda|\theta)\}\{\mathrm{Var.}(\lambda)\}^{-1}]^{1/2} .$$

The fact that the discrepancy of the mean of $\hat{\theta}^*_V$ for our five hundred hypothetical examinees from the expectation of θ is less than 0.001 supports this estimate for being a λ, the unbiased estimate of θ. Since the maximum likelihood estimate, $\hat{\theta}_V$, has a characteristic that for a fixed value of θ, it asymptotically distributes normally with θ and $[I(\theta)]^{-1/2}$ as the two parameters, as the amount of test information tends to infinity, and this convergence is relatively fast (Samejima, 1975, 1977a, 1977b), $E[\mathrm{Var.}(\hat{\theta}_V|\theta)]$ can be approximated by $E[I(\theta)^{-1}]$ for the interval, $(\underline{\theta},\bar{\theta})$. For Subtest 3, we find that

$$\text{(3.14)} \qquad E[I(\theta)^{-1}] \doteq 0.0803 .$$

From the above result we can see that the discrepancy of the variance of the modified maximum likelihood estimate, $\hat{\theta}^*_V$, for our sample from the population variance of θ to be 0.0457, which is less than $E[I(\theta)^{-1}]$ given as (3.14). If we substitute (3.14) for $E\{\mathrm{Var.}(\lambda|\theta)\}$ and $\hat{\theta}_V$ for λ in (3.12), we obtain for Subtest 3

$$\text{(3.15)} \qquad \mathrm{Var.}(\hat{\theta}_V) \doteq 2.1636 .$$

Substituting (3.14) and (3.15) into (3.13), we obtain for the product-moment correlation coefficient between θ and the maximum likelihood estimate $\hat{\theta}_V$,

$$(3.16) \qquad \mathrm{Corr.}(\theta, \hat{\theta}_V) \doteq 0.981 \quad .$$

It is interesting to note that our sample variance of the modified maximum likelihood estimate, $\hat{\theta}^*_V$, is slightly less than the estimated population variance of the maximum likelihood estimate, $\hat{\theta}_V$, and our sample correlation coefficient between θ and θ^*_V is slightly greater than the estimated population correlation coefficient between θ and the maximum likelihood estimate $\hat{\theta}_V$.

The error score, e_s , for each individual examinee s is defined by

$$(3.17) \qquad e_s = [\hat{\theta}^*_{V_s} - \theta_s][I(\theta_s)]^{-1/2} \quad ,$$

where θ_s is the ability level of the examinee s , and V_s indicates the response pattern obtained by the examinee s . Note in this definition of the error score the discrepancy of the estimated ability from the true ability is divided by the reciprocal of the square root of the amount of test information. Thus, if $\hat{\theta}^*_V$ distributes, approximately, normally with the true ability θ and the reciprocal of the square root of the test information function as the two parameters, then the error score e_s will conditionally distribute, approximately, normally with zero and unity as its two parameters, regardless of the ability level θ_s .

Figure 3-5 presents the cumulative frequency function of the error score e_s for the five hundred hypothetical examinees, which was obtained upon Subtest 3, together with the distribution function of the standard normal distribution. We can see in this figure that this cumulative frequency function is fairly close to the standard normal distribution function, the result which supports, though weakly, the approximate normality for the conditional distribution of e_s, given θ_s. The corresponding cumulative frequency function of e_s, which was obtained upon the original Old Test, is presented in Figure 3-6, together with the standard normal distribution function. Comparison of these two results reveals that the two cumulative frequency functions are very similar to each other, regardless of the fact that the former error score is defined for the modified maximum likelihood estimate $\hat{\theta}_V^*$, which includes twenty-six substitute estimates for negative and positive infinities, and the latter is with respect to the maximum likelihood estimate $\hat{\theta}_V$ itself, and that the square root of the test information function of Subtest 3 is unimodal while that of the original Old Test is constant ($\doteq 4.65$). The mean of the error score for Subtest 3 turned out to be -0.025 and the one for the original Old Test is -0.027, both of which are close to zero. The standard deviation, or the estimated second parameter, proved to be 0.972 for Subtest 3, and 0.995 for the original Old Test, in comparison with unity for the standard normal distribution. The variance of the error score is 0.944 for Subtest 3, and the one for the original Old Test is 0.990. It is interesting to note that

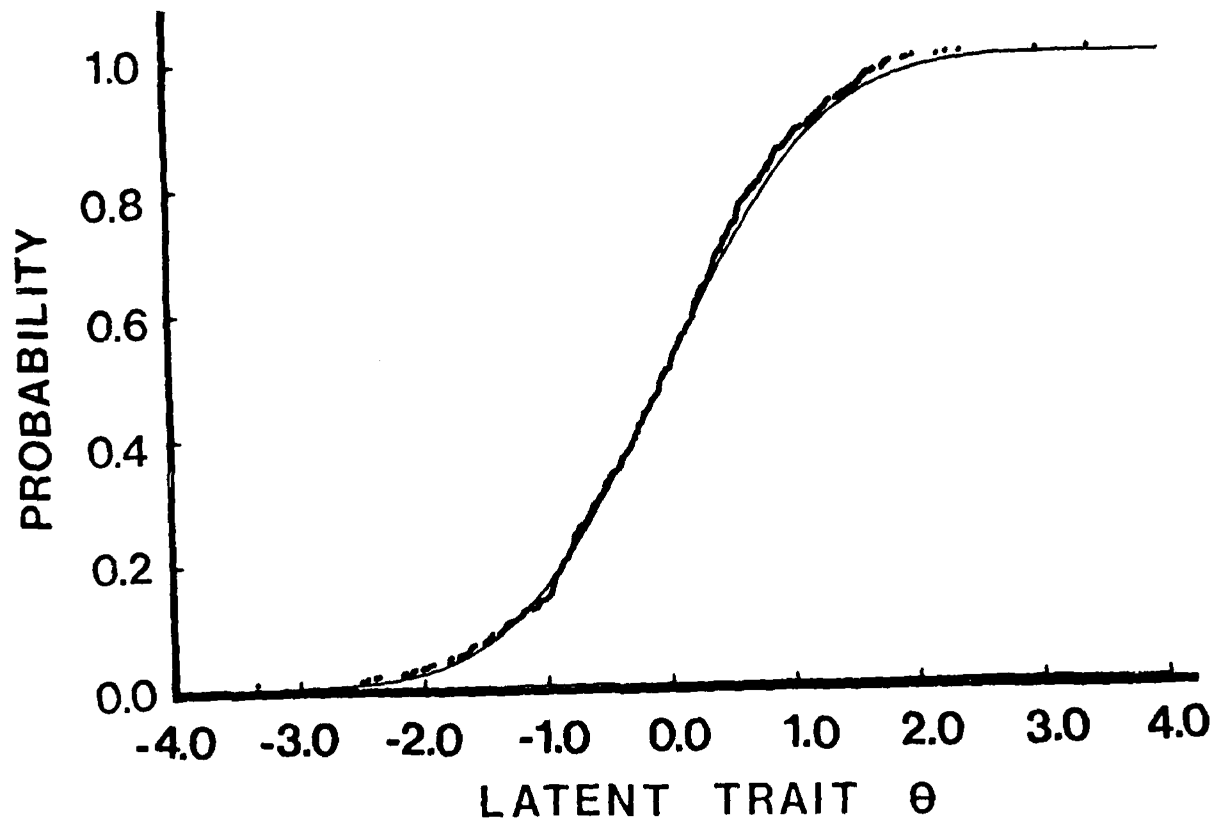


FIGURE 3-5

Cumulative Frequency Function of the Error Score e_s, Which Is Obtained upon Subtest 3, for the Five Hundred Hypothetical Examinees, Together with the Standard Normal Distribution Function.

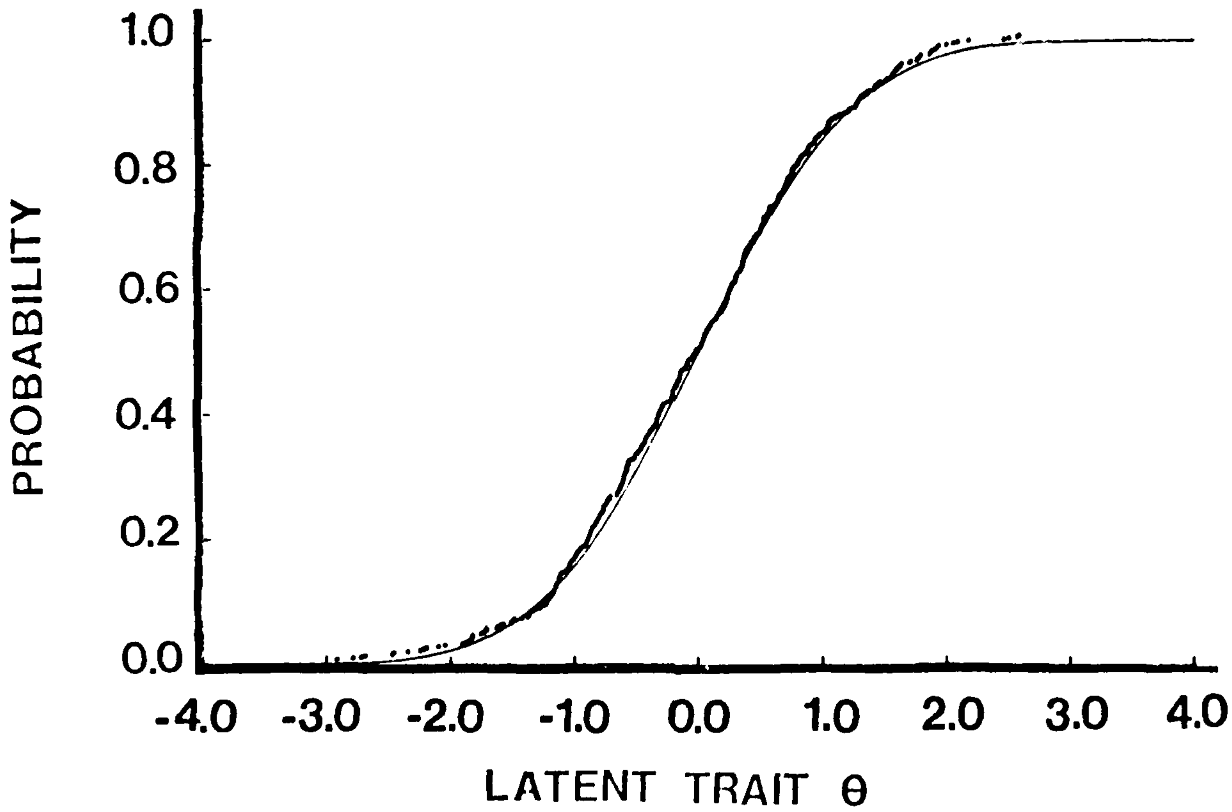


FIGURE 3-6

Cumulative Frequency Function of the Error Score e_s, Which Is Obtained upon the Original Old Test, for the Five Hundred Hypothetical Examinees, Together with the Standard Normal Distribution Function.

the error score for Subtest 3 has a less dispersion than the one for the original Old Test. This fact is inconsistent with the finding in the preceding chapter about the reduction of the standard error of estimation for θ^*_V compared with the the reciprocal of the square root of the test information function of LIS-U .

The sample linear regression of the error score e_s for Subtest 3 on ability θ turned out to be $0.003420\theta + 0.00009$, which is practically indistinguishable from the abscissa. This is even stronger support for the independence of the error score and ability θ than the result for the original Old Test, whose sample linear regression turned out to be $0.045790\theta + 0.00124$.

The frequency distribution of the five hundred error scores was obtained using the category width of 0.2 , for both Subtest 3 and the original Old Test. Figures 3-7 and 3-8 present these two results in the form of histogram, together with the standard normal density function. The chi-square test for the goodness of fit was made for each histogram against the standard normal density function, by combining all the categories below $e = -2.8$ and those above $e = 2.8$ into single categories, respectively. It turned out that $\chi^2_0 = 35.2252$ with 29 degrees of freedom, which provides us with $.20 < p < .30$, for Subtest 3, and $\chi^2_0 = 34.2248$ with 29 degrees of freedom, which also gives us $.20 < p < .30$, for the original Old Test, respectively. In this process of the chi-square test, there are fourteen categories for which the theoretical frequencies are less than ten. If we combine them appropriately

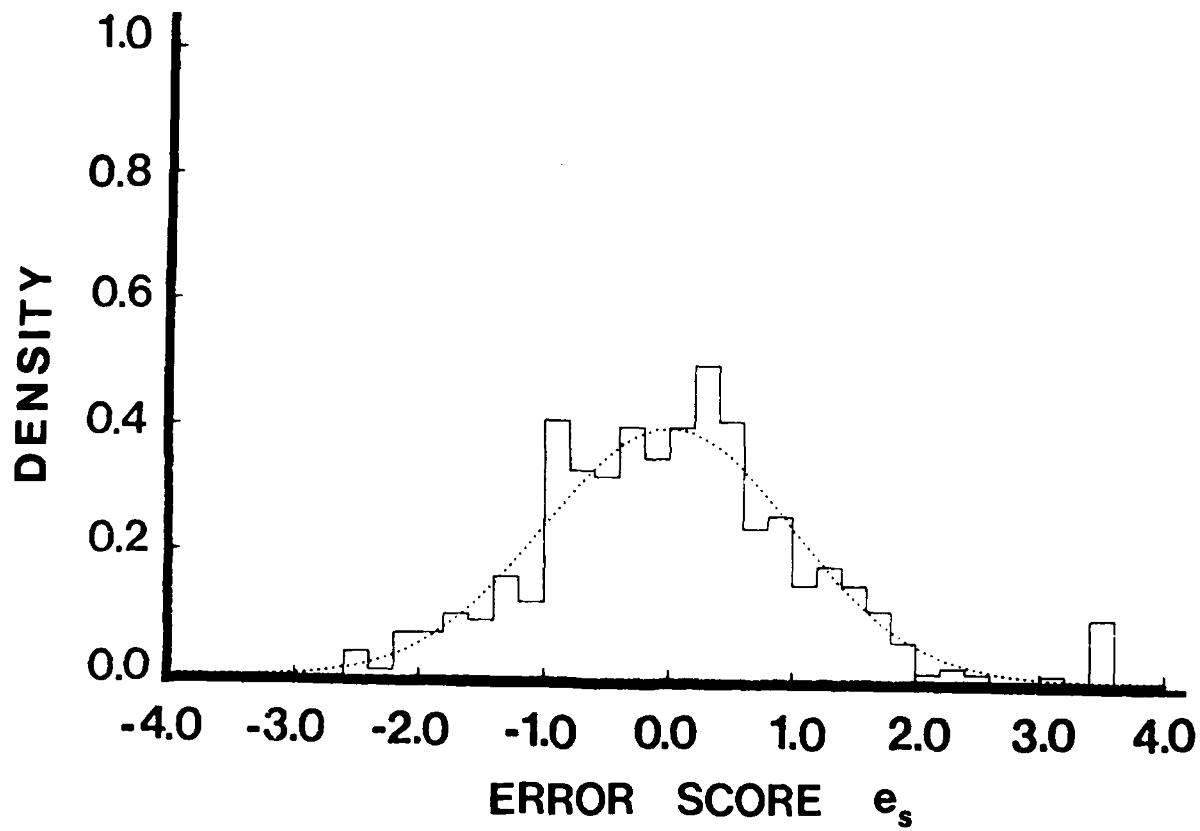


FIGURE 3-7

Frequency Distribution of the Error Score e_s , Which Is Based upon Subtest 3, for the Five Hundred Hypothetical Examinees, Compared with the Standard Normal Density Function.

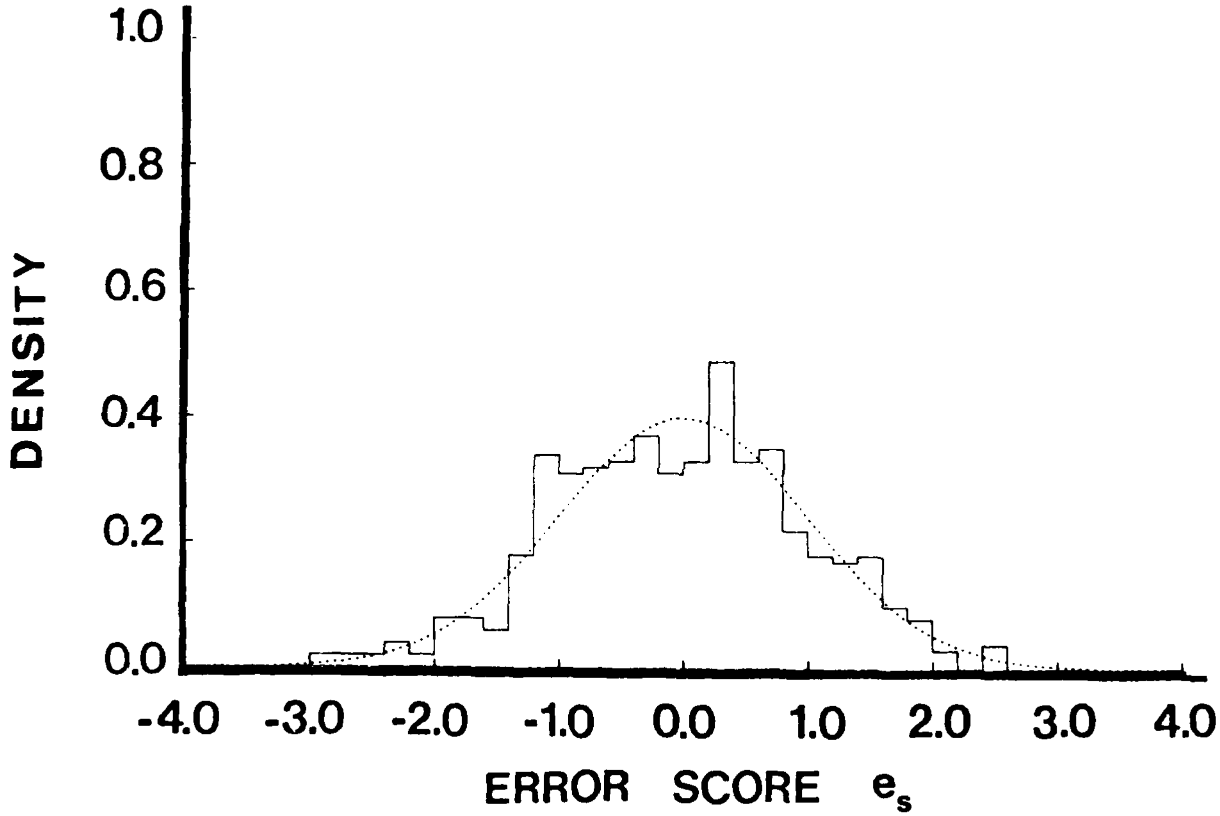


FIGURE 3-8

Frequency Distribution of the Error Score e_s, Which Is Based upon the *Original Old Test*, for the *Five Hundred Hypothetical* Examinees, Compared with the Standard Normal Density Function.

so that all the frequencies should become greater than, or equal to, ten, then the fits will be even better.

In the preceding chapter, we introduced and defined by (2.21) another population-free estimator, $\mu^{*'}_{1V}$. We notice that, although it is practically impossible to compute the values of this estimate when the number of possible response patterns is too large, as is the case with Subtest 3, we can compute these values for the two extreme response patterns, V-min and V-max , and used them as substitutes for negative and positive infinities of the maximum likelihood estimate.

The rationale behind these two estimates, $\mu^{*'}_{1V\text{-min}}$ and $\mu^{*'}_{V\text{-max}}$, may be given as follows. Let θ^{**}_V be an unknown estimator, which makes the integral of the conditional expectation of the squared error of estimation, given θ , for the interval, $(\underline{\theta},\bar{\theta})$, minimum. We define Q such that

$$(3.18) \qquad Q = \int_{\underline{\theta}}^{\bar{\theta}} E[(\theta^{**}_V - \theta)^2 | \theta] \; d\theta$$

$$= \sum_V \int_{\underline{\theta}}^{\bar{\theta}} (\theta^{**}_V - \theta)^2 \; P_V(\theta) \; d\theta \quad .$$

Differentiating (3.18) with respect to θ^{**}_V and setting the result equal to zero, we obtain

$$(3.19) \qquad \theta^{**}_V = \int_{\underline{\theta}}^{\bar{\theta}} \theta \; P_V(\theta) \; d\theta \; [\int_{\underline{\theta}}^{\bar{\theta}} P_V(\theta) \; d\theta]^{-1}$$

$$= \mu^{*'}_{1V} \quad .$$

Thus we have found that the estimator $\mu^{*'}_{1V}$ is the one which makes the integral of the conditional expectation of the squared error of estimation, given θ , for the interval, $(\underline{\theta}, \bar{\theta})$, minimum. Note that (3.19) includes no other response patterns, and is given as a function of V and the interval, $(\underline{\theta}, \bar{\theta})$ only. This implies that $\mu^{*'}_{1V}$ can be used as the estimate for a specific response pattern when those for any other response patterns are already given. For example, if we define θ^{***}_{V} in such a way that

$$(3.20) \qquad \theta^{***}_{V} \begin{cases} = \theta^{***}_{V\text{-min}} & \text{for } V = V\text{-min} \\ = \theta^{***}_{V\text{-max}} & \text{for } V = V\text{-max} \\ = \hat{\theta}_{V} & \text{otherwise ,} \end{cases}$$

and search for $\theta^{***}_{V\text{-min}}$ and $\theta^{***}_{V\text{-max}}$ following the above principle, then we will obtain

$$(3.21) \qquad \theta^{***}_{V\text{-min}} = \int_{\underline{\theta}}^{\bar{\theta}} \theta \, P_{V\text{-min}}(\theta) \, d\theta \left[\int_{\underline{\theta}}^{\bar{\theta}} P_{V\text{-min}}(\theta) \, d\theta \right]^{-1}$$

$$= \mu^{*'}_{1V\text{-min}} \quad ,$$

and

$$(3.22) \qquad \theta^{***}_{V\text{-max}} = \int_{\underline{\theta}}^{\bar{\theta}} \theta \, P_{V\text{-max}}(\theta) \, d\theta \left[\int_{\underline{\theta}}^{\bar{\theta}} P_{V\text{-max}}(\theta) \, d\theta \right]^{-1}$$

$$= \mu^{*'}_{1V\text{-max}} \quad .$$

This is exactly the case with the present situation, which provides us with the justification for using $\mu^{*'}_{1V\text{-min}}$ and $\mu^{*'}_{1V\text{-max}}$ for the two extreme response patterns, V-min and V-max , which substitute

for the negative and positive infinities, respectively, of the maximum likelihood estimate. An advantage of these two estimates, $\mu^{*\prime}_{1V-min}$ and $\mu^{*\prime}_{1V-max}$, over $\hat{\theta}^{*}_{V-min}$ and $\hat{\theta}^{*}_{V-max}$, is that they are theoretical values, and obtainable without depending upon the Monte Carlo method. Note, however, that no consideration for the conditional unbiasedness of the estimator is given in adopting $\mu^{*\prime}_{1V-min}$ and $\mu^{*\prime}_{1V-max}$. We computed these values for Subtest 3 using the interval of θ, $(-2.5, 2.5)$, and they turned out to be -2.2684 and 2.2884, respectively. Comparison of these values with $\hat{\theta}^{*}_{V-min}$ $(= -2.3145)$ and $\hat{\theta}^{*}_{V-max}$ $(= 2.0403)$ reveals that they are not too far from each other.

It should be noted that the above values of $\hat{\theta}^{*}_{V-min}$ and $\hat{\theta}^{*}_{V-max}$ are not the least and greatest values of $\hat{\theta}^{*}_{V}$. In fact, among the four hundred and seventy-four finite values of the maximum likelihood estimate, we find such values as -2.4698 (8), -2.3887 (5), -2.3846 (2) and -2.3585 (6), which are less than -2.3145, and 2.4651 (7), 2.3526 (12), 2.3454 (7), 2.3359 (2), 2.2762 (3), 2.0885 (5) and 2.0789 (3), which are greater than 2.0403, with the integer attached to each value in parenthesis indicating the corresponding frequency. Considering that fact that $\hat{\theta}^{*}_{V-min}$ and $\hat{\theta}^{*}_{V-max}$ substitute for negative and positive infinities of the maximum likelihood estimate, it may be hard to accept these values. This fact indicates that the effectiveness of Subtest 3 as an instrument for measuring ability θ extends itself for a greater interval of θ than $(-2.5, 2.5)$. From the definition of the

substitute estimates, $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$, and also from the findings in the preceding chapter with respect to $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, we can expect to be able to find an optimal interval of θ for which the modified maximum likelihood estimate, $\hat{\theta}^*_V$, is, approximately, conditionally unbiased, with $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$ being the least and the greatest values of the estimate $\hat{\theta}^*_V$.

IV Modified Maximum Likelihood Estimate for the Transformed Latent Trait

We have seen in the preceding chapter that, for Subtest 3, the two alternative estimates, $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$, for the two extreme response patterns may be too small in absolute values, if we take the interval of θ, $(-2.5, 2.5)$, for $(\underline{\theta},\bar{\theta})$, to accept as the substitutes for the negative and positive infinities of the maximum likelihood estimate. The logical step we should take next will, therefore, be to search for an optimal interval for $(\underline{\theta},\bar{\theta})$ for this purpose. This can be done by expanding the interval as far as possible in both negative and positive directions, with the restriction that the resultant $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$ provide us with an, approximately, conditionally unbiased estimator $\hat{\theta}^*_V$ for that interval of θ.

We notice, however, that we need a transformation of θ to τ in order to use a test like Subtest 3, which does not have a constant test information function, as the Old Test in our methods of estimating the operating characteristics of discrete item responses (Samejima, RR-80-2, RR-80-4). The search for the alternative estimates for the two extreme response patterns, V-min and V-max, will, therefore, become more meaningful if we do it with respect to τ, which provides us with a constant test information function, $I^*(\tau)$, for Subtest 3.

Let $P^*_V(\tau)$ be the operating characteristic of the response pattern V, which is defined as a function of the transformed

latent trait τ . This conditional probability, given ability θ , stays the same as the original operating characteristic, $P_V(\theta)$, as far as τ is a strictly increasing, or decreasing, function of θ . Thus we can write

$$(4.1) \qquad P^*_V(\tau) = P_V(\theta) \quad .$$

Let $\tau^*_{V\text{-min}}$ and $\tau^*_{V\text{-max}}$ denote the estimates of τ which are analoguous to $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$ defined by (2.14) .
We can write

$$(4.2) \quad \begin{cases} \tau^*_{V\text{-min}} = [\frac{1}{2}(\tau_c^2 - \underline{\tau}^2) - \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\tau}_V \int_{\underline{\tau}}^{\tau_c} P^*_V(\tau)\, d\tau] \\ \qquad\qquad [\int_{\underline{\tau}}^{\tau_c} P^*_{V\text{-min}}(\tau)\, d\tau]^{-1} \\ \tau^*_{V\text{-max}} = [\frac{1}{2}(\bar{\tau}^2 - \tau_c^2) - \sum\limits_{\substack{V\neq V\text{-min}\\ V\neq V\text{-max}}} \hat{\tau}_V \int_{\tau_c}^{\bar{\tau}} P^*_V(\tau)\, d\tau] \\ \qquad\qquad [\int_{\tau_c}^{\bar{\tau}} P^*_{V\text{-max}}(\tau)\, d\tau]^{-1} \quad , \end{cases}$$

where $\underline{\tau}$ and $\bar{\tau}$ indicate the lower and upper endpoints of an appropriately selected interval of τ , τ_c is a critical value of τ below which $P^*_{V\text{-max}}(\tau)$ assumes negligibly small values, and above which so does $P^*_{V\text{-min}}(\tau)$, and $\hat{\tau}_V$ is the maximum likelihood estimate of τ which is assigned to the response pattern V . Let us assume that the first three values, $\underline{\tau}$, $\bar{\tau}$ and τ_c , are

directly transformed from $\underline{\theta}$, $\bar{\theta}$ and θ_c , respectively, through the strictly increasing transformation

$$(4.3) \qquad \tau = \tau(\theta) = \sum_{k=0}^{8} \alpha_k^* \theta^k \qquad \text{for} \quad -4.0 \leqslant \theta \leqslant 4.0 ,$$

whose nine coefficients are given in Table 4-2 . The critical value, τ_c , which was transformed from θ_c (= -0.4146) through (4.3), turned out to be -0.5455 . Again, this value of τ does <u>not</u> make the product of the two operating characteristics, $P^*_{V\text{-min}}(\tau)$ and $P^*_{V\text{-max}}(\tau)$, maximal, but is deviated toward the negative side. It should be noted that the maximum likelihood estimate, $\hat{\tau}_V$, of the transformed ability τ is also given as the direct transformation of $\hat{\theta}_V$ (Samejima, 1969) through (4.3), for every response pattern V . Thus we can rewrite (4.2) in the form,

$$(4.4) \quad \begin{cases} \tau^*_{V\text{-min}} = [\frac{1}{2}\{\tau(\theta_c)^2 - \tau(\underline{\theta})^2\} - \sum_{\substack{V\neq V\text{-min} \\ V\neq V\text{-max}}} \tau(\hat{\theta}_V)\int_{\underline{\theta}}^{\theta_c} P_V(\theta)\,\frac{d\tau}{d\theta}\,d\theta \\ \qquad\qquad [\int_{\underline{\theta}}^{\theta_c} P_{V\text{-min}}(\theta)\,\frac{d\tau}{d\theta}\,d\theta\,]^{-1} \\ \tau^*_{V\text{-max}} = [\frac{1}{2}\{\tau(\bar{\theta})^2 - \tau(\theta_c)^2\} - \sum_{\substack{V\neq V\text{-min} \\ V\neq V\text{-max}}} \tau(\hat{\theta}_V)\int_{\theta_c}^{\bar{\theta}} P_V(\theta)\,\frac{d\tau}{d\theta}\,d\theta \\ \qquad\qquad [\int_{\theta_c}^{\bar{\theta}} P_{V\text{-max}}(\theta)\,\frac{d\tau}{d\theta}\,d\theta\,]^{-1} \end{cases} .$$

It is obvious from (4.4) that, in general, we have

$$\text{(4.5)} \qquad \begin{cases} \tau^*_{V\text{-min}} \neq \tau(\theta^*_{V\text{-min}}) \\ \tau^*_{V\text{-max}} \neq \tau(\theta^*_{V\text{-max}}) \end{cases} .$$

The sample statistic versions of $\tau^*_{V\text{-min}}$ and $\tau^*_{V\text{-max}}$, which will be denoted by $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$, respectively, are defined by

$$\text{(4.6)} \qquad \begin{cases} \hat{\tau}^*_{V\text{-min}} = [\frac{1}{2}(\tau_c + \underline{\tau})\ N_L - \sum\limits_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\tau}_V\ N_{LV}]\ N^{-1}_{LV\text{-min}} \\ \hat{\tau}^*_{V\text{-max}} = [\frac{1}{2}(\underline{\tau} + \tau_c)\ N_H - \sum\limits_{\substack{V \neq V\text{-min} \\ V \neq V\text{-max}}} \hat{\tau}_V\ N_{HV}]\ N^{-1}_{HV\text{-max}} \end{cases} ,$$

where N_L, N_H, N_{LV}, N_{HV}, $N_{LV\text{-min}}$ and $N_{HV\text{-max}}$ are as defined in the preceding chapter. From (4.6) and (3.3), it is obvious that, again, in general, we have

$$\text{(4.7)} \qquad \begin{cases} \hat{\tau}^*_{V\text{-min}} \neq \tau(\hat{\theta}^*_{V\text{-min}}) \\ \hat{\tau}^*_{V\text{-max}} \neq \tau(\hat{\theta}^*_{V\text{-max}}) \end{cases} .$$

We must make our choice, therefore, as to which of the two sets of the alternative estimates should be taken. In this chapter, our choice is to take $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$.

The transformation of θ to τ for Subtest 3 starts from the approximation of the square root of the test information function by a polynomial, using the method of moments (Elderton and Johnson, 1969, Johnson and Kotz, 1970), as was done for some other subtests of the original Old Test in a previous study (Samejima, RR-80-4). Note that such a polynomial is the best fitted polynomial of a given

degree in the least squares principle (Samejima and Livingston, RR-79-2).

The degree of the polynomial selected here is seven, and the interval of θ for which the method of moments was applied is $(-4.0, 4.0)$. The eight coefficients of the resultant polynomial, $\sum_{k=0}^{7} \alpha_k \theta^k$, are presented in Table 4-1 in the ascending order of k, and the polynomial is shown by a dotted line in Figure 3-1 of the preceding chapter, together with the original square root of the test information function of Subtest 3. We can see in this figure that the polynomial thus obtained provides us with an extremely good approximation to the square root of the test information function of Subtest 3. Using this approximated polynomial, the transformation of ability θ to τ is also given in the form of another polynomial (Samejima, RR-80-2), such that

$$\tau(\theta) = \sum_{k=0}^{8} \alpha_k^* \theta^k \quad , \tag{4.8}$$

where

$$\alpha_k^* \begin{cases} = d & k = 0 \\ = (CK)^{-1} \alpha_{k-1} & k = 1,2,\ldots,8 \end{cases} \quad , \tag{4.9}$$

with d indicating an arbitrarily set constant, and C being the constant which the square root of the test information function, $[I^*(\tau)]^{1/2}$, assumes for the interval of τ of our interest. In the present study, we use $d = 0$ and $C = 3.5$. The coefficients, α_k^*, of the resultant polynomial, which transforms θ to τ, are

TABLE 4-1

Coefficients of the Polynomial of Degree 7 Obtained by the Method of Moments Using the Interval of θ , (-4.0, 4.0) , to Approximate the Square Root of the Test Information Function of Subtest 3.

k	α_k
0	0.46408884D+01
1	0.60789659D-01
2	-0.41482735D+00
3	0.14684659D-01
4	0.51686862D-02
5	-0.36903316D-02
6	0.21313602D-03
7	0.15726020D-03

shown in the ascending order of k in Table 4-2, and the functional relationship between θ and τ is observed in Figure 4-1. The two operating characteristics, $P^*_{V\text{-}min}(\tau)$ and $P^*_{V\text{-}max}(\tau)$, are shown by solid and dotted lines, respectively, together with the position of the critical value, τ_c (= -0.5455), in Figure 4-2.

As for the interval, $(\underline{\tau}, \bar{\tau})$, seven different cases were chosen more or less arbitrarily, and are shown as Cases 1 through 7 in Table 4-3. The intervals were selected in such a way that we set the values of $\min\{\sqrt{I(\theta)}\}$, the lower bound of the square root of the test information function, and then corresponding intervals, $(\underline{\theta}, \bar{\theta})$, were determined, and, finally, the pairs of values, $\underline{\tau}$ and $\bar{\tau}$, were obtained through (4.8). In addition to these seven cases, another interval of τ, (-3.0, 3.0), was added as Case 8.

The number of hypothetical examinees for each of the eight cases was determined in the following way. It was intended that these numbers should be substantially larger than five hundred, which was used in the preceding chapter, in order to decrease the error caused by the Monte Carlo method. For Case 8, we use five thousand hypothetical examinees, or $N = 5{,}000$. They were placed at the one thousand positions of τ, which start from -2.997 and end with 2.997, with equal steps of 0.006, with five hypothetical examinees sharing each position. For Case 7, out of these five thousand hypothetical examinees, those who were located outside of the interval, (-2.8267, 2.8095), were excluded. Thus the total number of the hypothetical examinees is 4,695 in Case 7,

TABLE 4-2

Coefficients of the Polynomial of Degree 8 to Transform θ to τ for Subtest 3.

k	α_k^*
0	0.00000000D 00
1	0.13259652D 01
2	0.86842420D-02
3	-0.39506409D-01
4	0.10489276D-02
5	0.29536370D-03
6	-0.17572918D-03
7	0.86989735D-05
8	0.56164139D-05

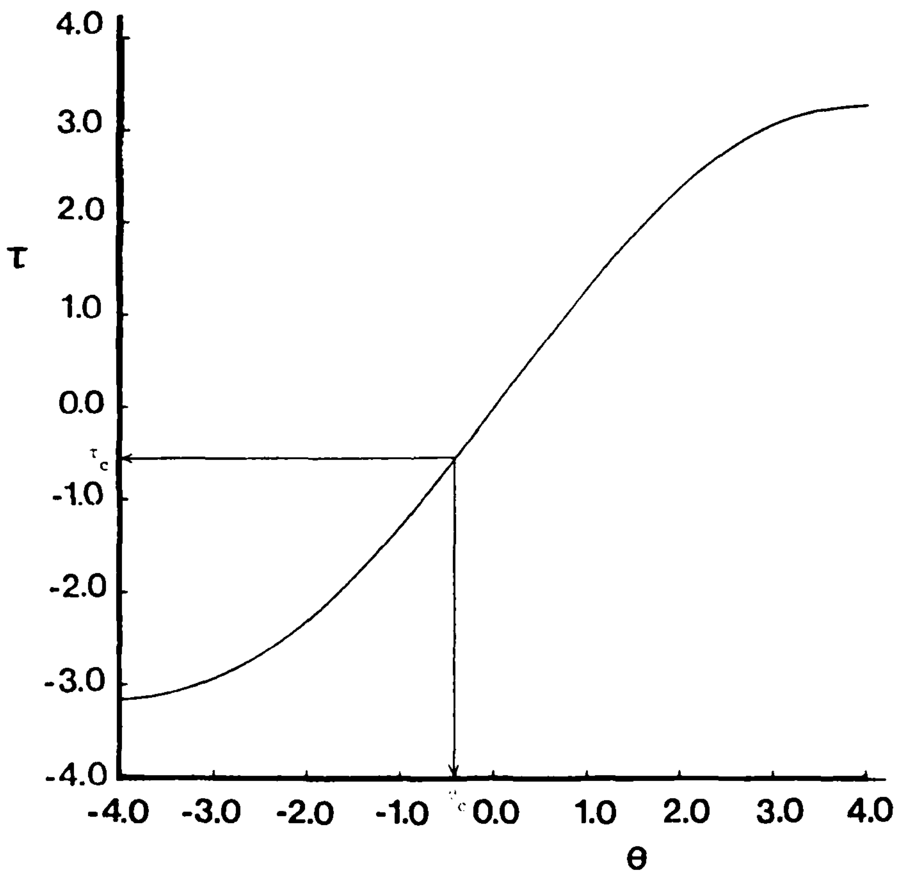


FIGURE 4-1

Functional Relationship Between θ and τ Based upon Subtest 3. The Positions of θ_c and τ_c Are Also Indicated.

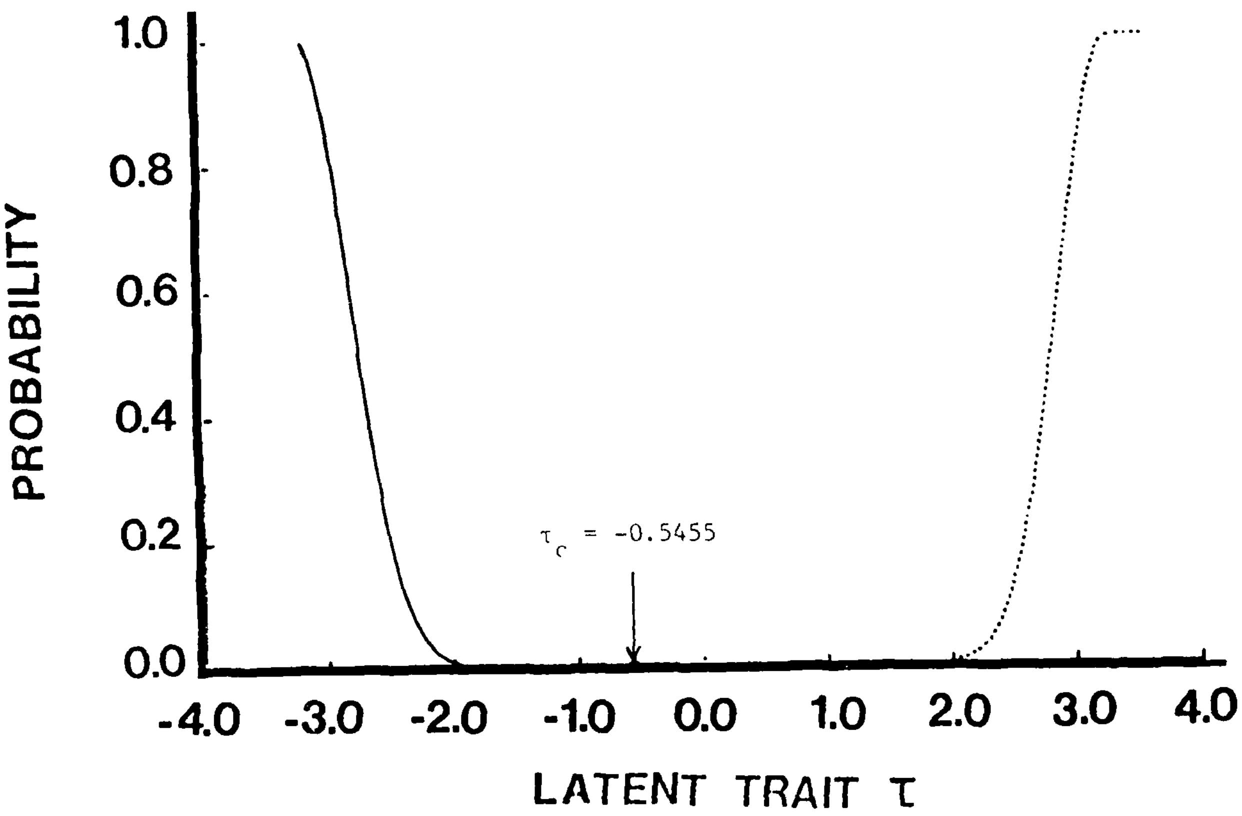


FIGURE 4-2

Operating Characteristics of V-min (Solid Line) and V-max (Dotted Line) Given As Functions of the Transformed Latent Trait τ , Together with the Position of the Critical Value, τ_c .

with the exclusion of the first 145 examinees and the last 160 examinees from the original five thousand. In each of the remaining six cases, the number of hypothetical examinees was determined in the similar manner, and it is presented in the last column of Table 4-3. From this total number of examinees, N, the two sample sizes, N_L and N_H, were determined in each case, depending upon how many examinees were positioned below and above the critical value, τ_c (= -0.5455). These numbers are also presented in Table 4-3. Thus, in each case, these hypothetical examinees can be considered as a sample from the uniform distribution for the interval, $(\underline{\tau},\bar{\tau})$. Note, however, that, because of the way the examinees were selected, the values $\underline{\tau}$ and $\bar{\tau}$ were slightly shifted for Cases 1 through 7. The new endpoints of the interval of τ are presented in Table 4-5 for the four cases, Cases 4 through 7.

It turned out that for the first three cases, Cases 1 through 3, the two frequencies, $N_{LV\text{-}min}$ and $N_{HV\text{-}max}$, are so small, i.e., 1 and 3 for Case 1, 1 and 10 for Case 2, and 8 and 19 for Case 3, respectively. This is due to the fact that these three intervals of τ are relatively small, and the probability with which the examinee, whose ability level is within each interval, obtains either V-min or V-max is low. With these small frequencies substituted in (4.6), we obtained such absurd results for $\hat{\tau}^*_{V\text{-}min}$ and $\hat{\tau}^*_{V\text{-}max}$ as 7.7998 and -2.2507 for Case 1, 11.3745 and 0.1132 for Case 2, and -0.8183 and 1.4841 for Case 3, respectively. It is obvious that we should not take these results

TABLE 4-3

Lower Bound of the Square Root of the Test Information Function, $\text{Min}\{\sqrt{I(\theta)}\}$, of Subtest 3, the Two Endpoints of the Interval, $\underline{\theta}$ and $\bar{\theta}$, the Transformed Endpoints, $\underline{\tau}$ and $\bar{\tau}$, the Numbers of Hypothetical Subjects, N_L and N_H , for the Two Intervals, $(\underline{\tau}, \tau_c)$ and $(\tau_c, \bar{\tau})$, and the Total Number of Subjects , N , in Each of the Eight Cases.

Case	$\min\{\sqrt{I(\theta)}\}$	$\underline{\theta}$	$\bar{\theta}$	$\underline{\tau}$	$\bar{\tau}$	N_L	N_H	N
1	3.50	-1.5	1.7	-1.8456	2.0771	1,085	2,185	3,270
2	3.25	-1.7	1.9	-2.0521	2.2668	1,255	2,345	3,600
3	3.00	-1.9	2.1	-2.2461	2.4373	1,415	2,485	3,900
4	2.75	-2.1	2.3	-2.4273	2.5860	1,570	2,610	4,180
5	2.50	-2.2	2.4	-2.5131	2.6516	1,640	2,665	4,305
6	2.25	-2.4	2.6	-2.6757	2.7636	1,775	2,760	4,535
7	2.00	-2.6	2.7	-2.8267	2.8095	1,900	2,795	4,695
8	---	---	---	-3.0000	3.0000	2,045	2,955	5,000

seriously, and we must conclude that these three intervals of τ are too small for our purpose of obtaining the two estimates, $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$. In all eight cases, both $N_{LV\text{-max}}$ and $N_{HV\text{-max}}$ turned out to be zero, the fact that indicated the success in selecting the critical value, τ_c .

Table 4-4 presents the resultant values of $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$, together with the two frequencies, $N_{V\text{-min}}$ and $N_{V\text{-max}}$, for each of Cases 4 through 8 . These two estimates increase in absolute values as the interval becomes larger, as is expected from their definitions.

The sample regressions of $\hat{\tau}^*_V$ on τ for Cases 4 through 8 are presented in Figures 4-3 through 4-7, respectively. In each graph, the mean of five $\hat{\tau}^*_V$'s corresponding to a fixed value of τ is plotted as one point, to make the total number of points 836 for Case 4 , 861 for Case 5 , 907 for Case 6 , 939 for Case 7 , and 1,000 for Case 8 , respectively. We can see that, in each case, these points of sample regression cluster around the unbiasedness line, or the straight line with forty-five degrees from the abscissa passing the origin, (0,0) , which is shown in each of the five figures.

Table 4-5 presents the sample mean and variance of τ , those of $\hat{\tau}^*_V$, and the sample product-moment correlation coefficient of τ and $\hat{\tau}^*_V$, together with the two endpoints of the interval, $(\underline{\tau},\bar{\tau})$, for each of Cases 4 through 8 . In the same table, also presented in parentheses are the corresponding expectations, population variances and population correlation coefficients. These values

TABLE 4-4

Two Estimates, $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$, and the Numbers of Hypothetical Subjects, $N_{V\text{-min}}$ and $N_{V\text{-max}}$, Who Obtained Either of the Two Extreme Response Patterns, V-min and V-max, Respectively, in Each of the Five Cases, Cases 4 through 8.

Case	$\hat{\tau}^*_{V\text{-min}}$	$N_{V\text{-min}}$	$\hat{\tau}^*_{V\text{-max}}$	$N_{V\text{-max}}$
4	-1.6061	23	2.0856	32
5	-2.0651	39	2.2750	42
6	-2.4788	81	2.5455	74
7	-2.6867	145	2.6865	93
8	-2.8214	258	2.8596	196

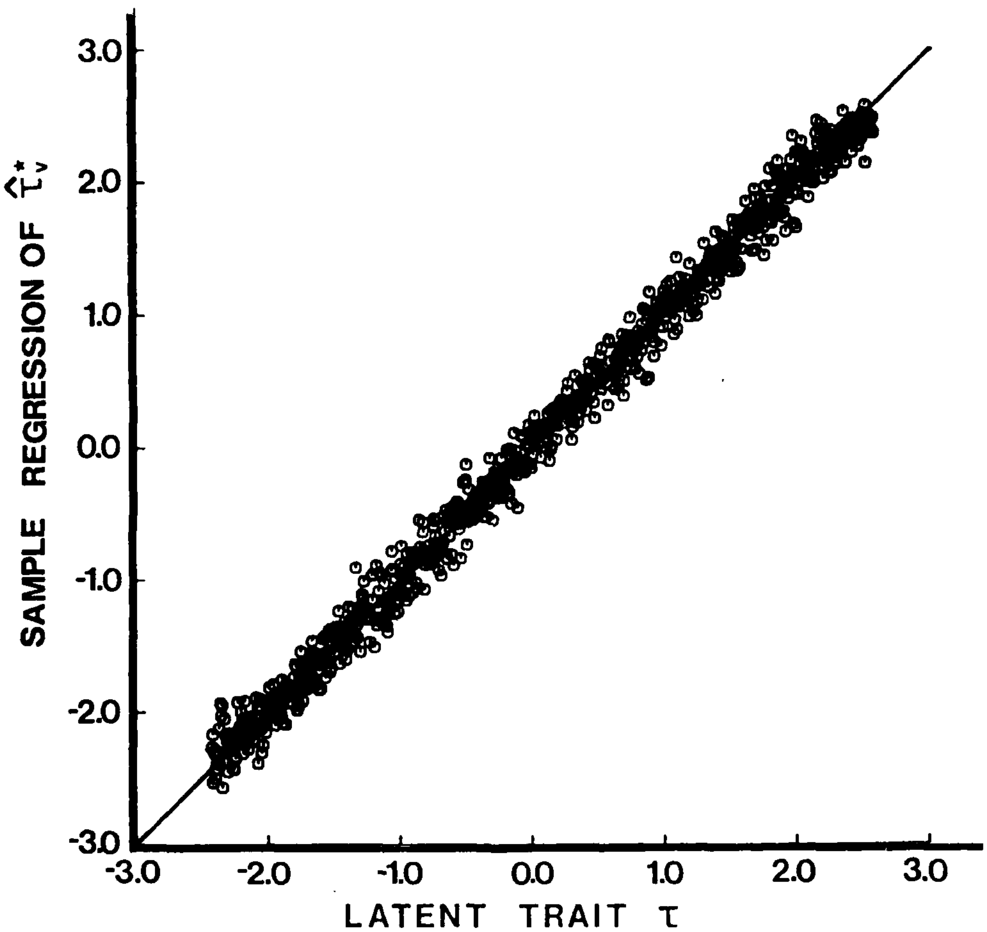


FIGURE 4-3

Sample Regression of $\hat{\tau}_V^*$ on τ, for 836 Fixed Values of τ.

Case 4

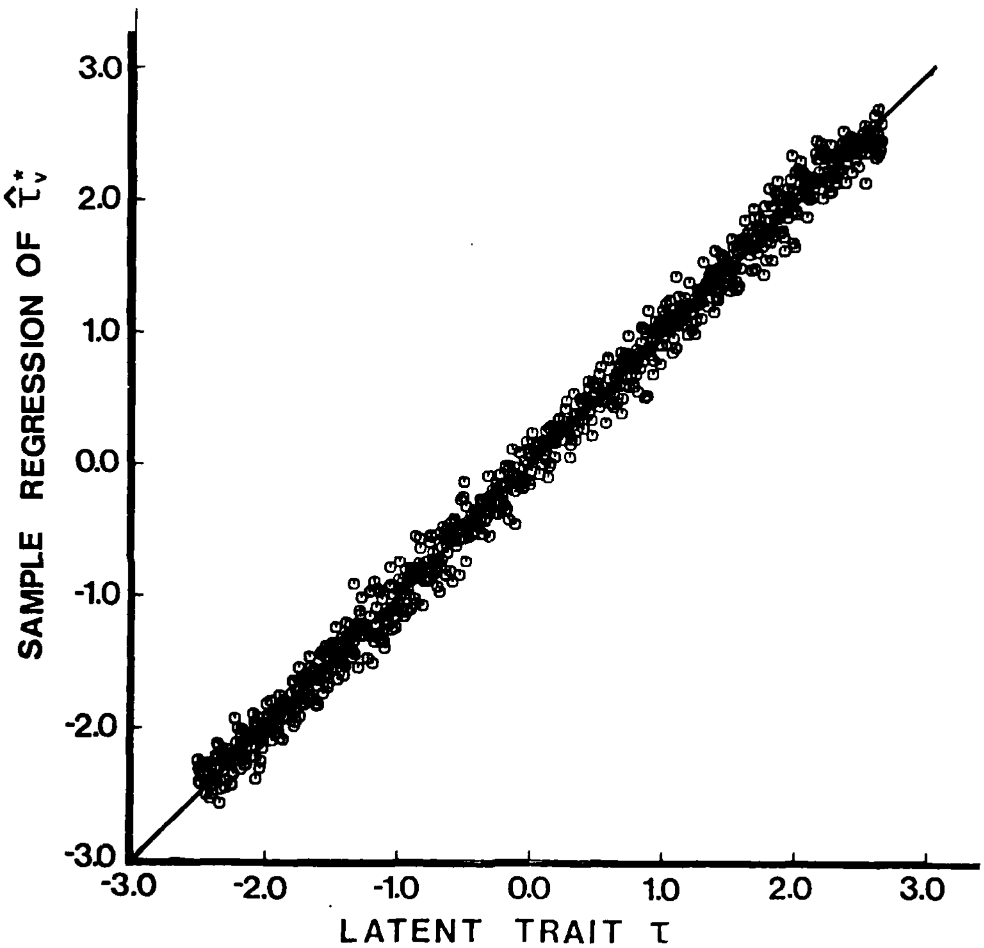


FIGURE 4-4

Sample Regression of $\hat{\tau}_V^*$ on τ, for 861 Fixed Values of τ.

Case 5

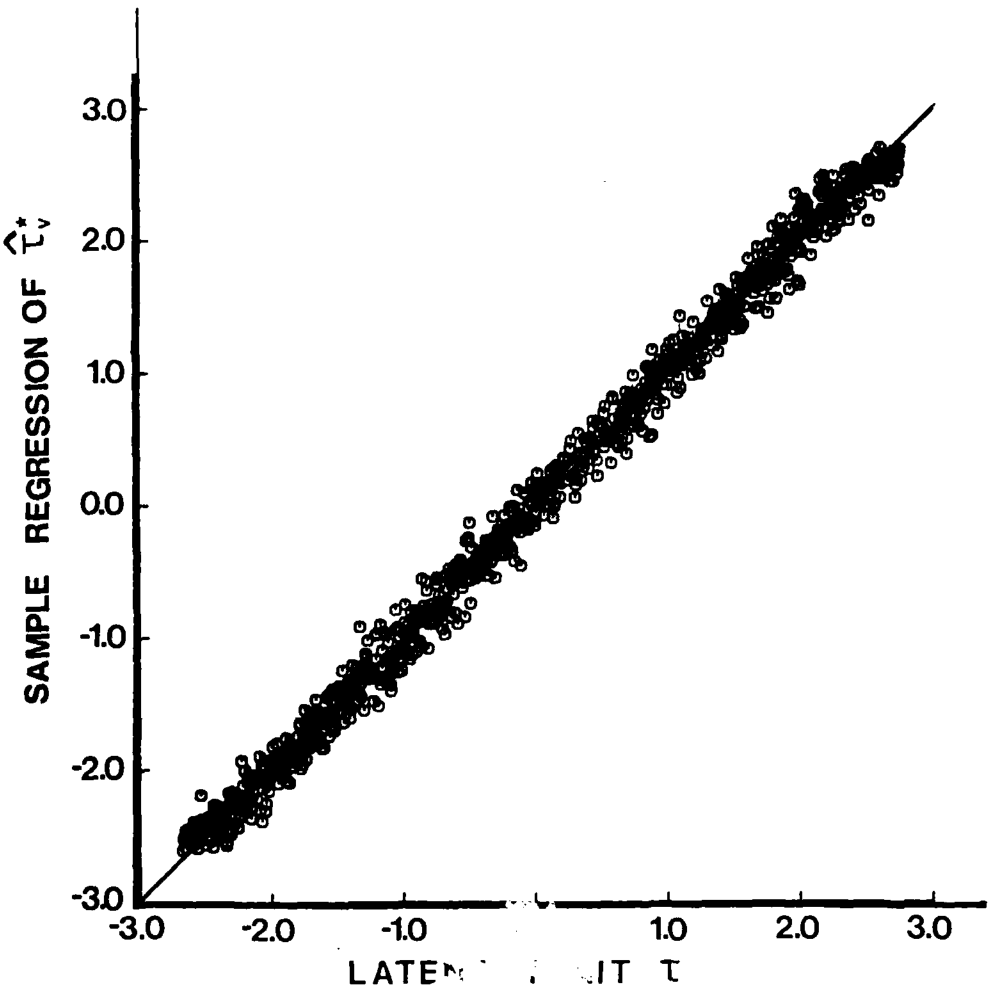


FIGURE 4-5

Sample Regression of $\hat{\tau}_V^*$ on τ , for 907 Fixed Values of τ .
Case 6

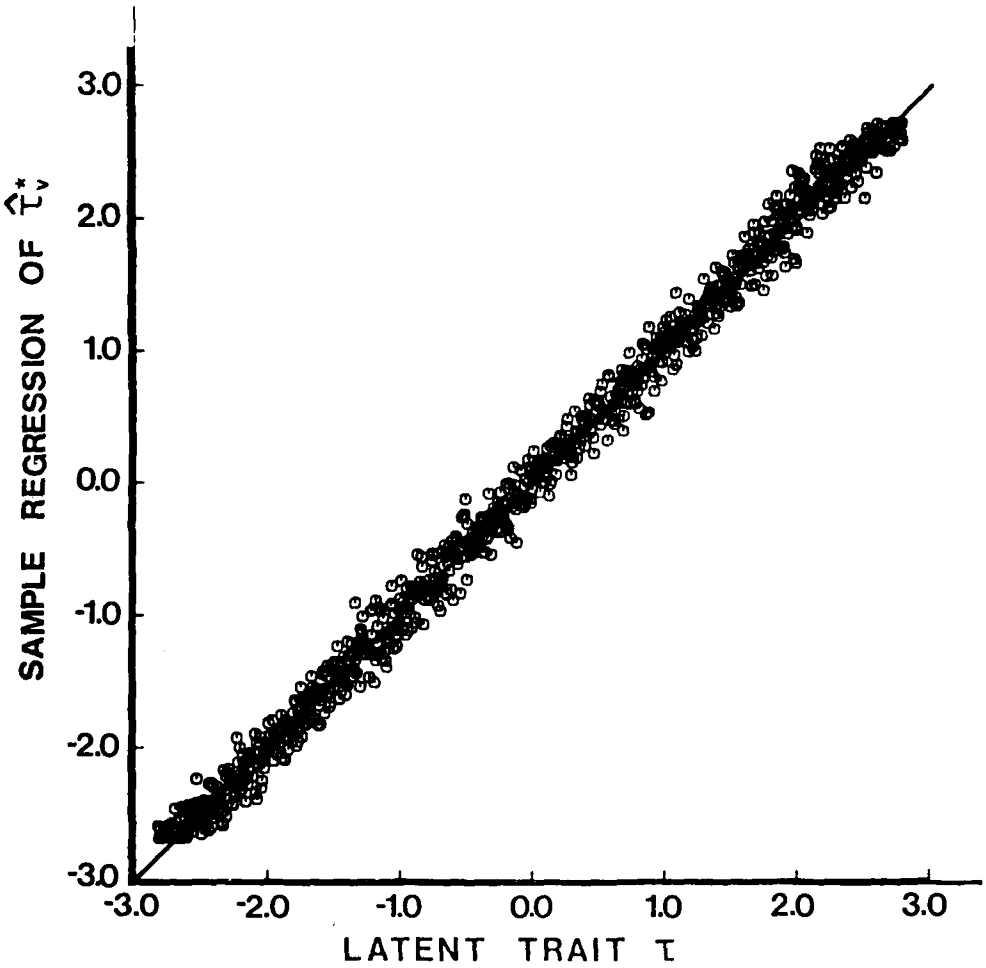


FIGURE 4-6

Sample Regression of $\hat{\tau}_V^*$ on τ, for 939 Fixed Values of τ.

Case 7

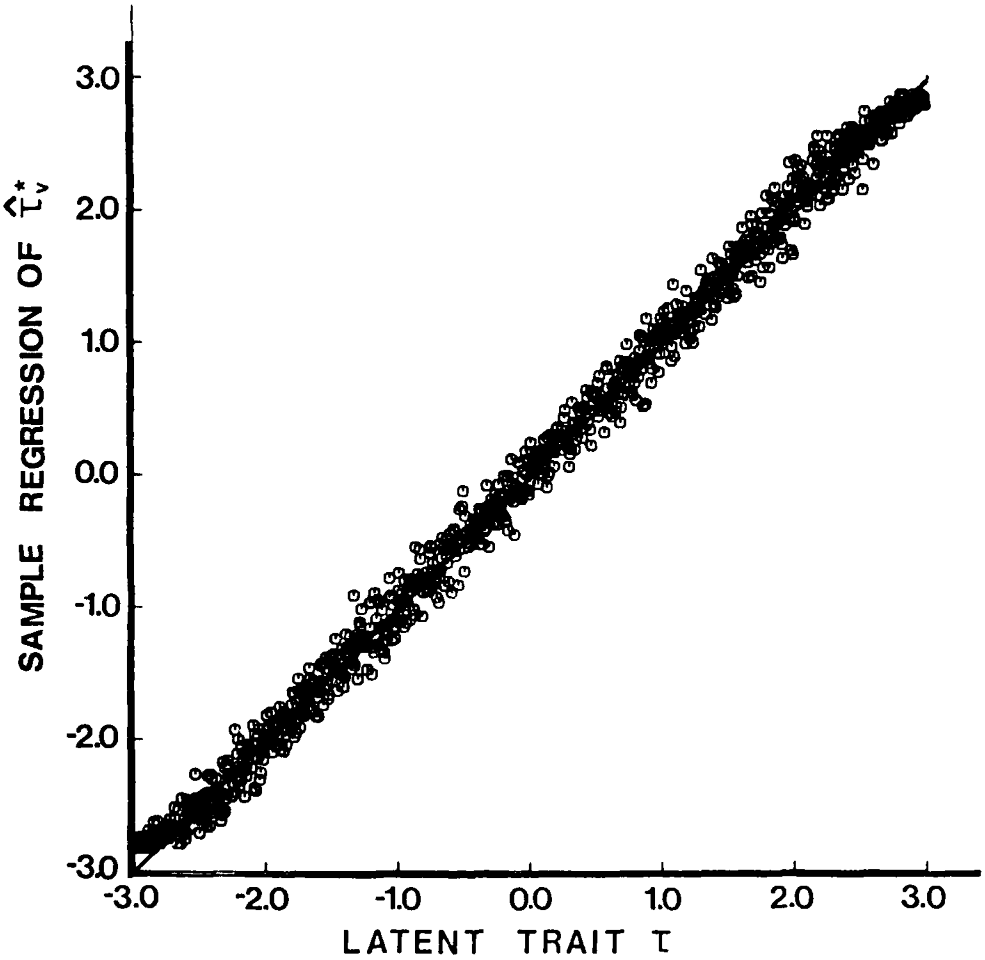


FIGURE 4-7

Sample Regression of $\hat{\tau}_V^*$ on τ, for 1,000 Fixed Values of τ.
Case 8

TABLE 4-5

Mean and Variance of τ, Those of $\hat{\tau}^*_V$, the Product-Moment Correlation Coefficient of τ and $\hat{\tau}^*_V$, and the Two Endpoints of the Interval, $(\underline{\tau},\bar{\tau})$, in Each of the Five Cases.

Case	τ mean	τ variance	$\hat{\tau}^*_V$ mean	$\hat{\tau}^*_V$ variance	Corr.$(\tau,\hat{\tau}^*_V)$	$\underline{\tau}$	$\bar{\tau}$
4	0.07800 (0.078)	2.09668 (2.09669)	0.07879 (0.078)	2.16160 (2.17832)	0.98081 (0.98108)	-2.430	2.586
5	0.06900 (0.069)	2.22396 (2.22396)	0.06929 (0.069)	2.28554 (2.30560)	0.98254 (0.98214)	-2.514	2.652
6	0.04500 (0.045)	2.46794 (2.46795)	0.04458 (0.045)	2.52176 (2.54958)	0.98475 (0.98386)	-2.676	2.766
7	-0.00900 (-0.009)	2.64516 (2.64516)	-0.00844 (-0.009)	2.70365 (2.72680)	0.98589 (0.98492)	-2.826	2.808
8	0.00000 (0.000)	3.00000 (3.0000)	0.00027 (0.000)	3.05110 (3.08163)	0.98762 (0.98667)	-3.000	3.000

were obtained by replacing τ for θ in (3.6), (3.7), (3.10), (3.11), (3.12) and (3.13), and replacing $\hat{\tau}^*_V$ for λ in the last four of these six formulas, and by using C^{-2} ($\doteq$ 0.081633) for $E\{Var.(\hat{\tau}^*_V|\tau)\}$. We can see that, in each case, these sample means, variances and correlation coefficients are very close to the corresponding population parameters. It is interesting to note, however, that there is a mild tendency that the sample variance of $\hat{\tau}^*_V$ is less than the population variance, and the sample correlation coefficient between τ and $\hat{\tau}^*_V$ is greater than the population correlation coefficient.

Table 4-6 presents the two coefficients of the linear regression, $\alpha\tau + \beta$, of $\hat{\tau}^*_V$ on τ , or the best fitted line in the least squares principle, for each of Cases 4 through 8 . We can see that the first coefficient, α , is very close to unity, and the second coefficient, β , is very close to zero, in each of the five cases, and the linear regression is practically the same as the unbiasedness line, or the line with forty-five degrees from the abscissa passing the origin, (0,0) . Evidently, the two alternative estimates, $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$, turned out to be suitable substitutes for the negative and positive infinities of the maximum likelihood estimate so that the resultant $\hat{\tau}^*_V$ be, approximately, conditionally unbiased for the interval, $(\underline{\tau},\bar{\tau})$, in each of the five cases. If we extend the interval beyond $(\underline{\tau},\bar{\tau})$, however, the approximate unbiasedness of $\hat{\tau}^*_V$ does not necessarily hold. The expansion of the interval to (-3.0, 3.0) for Case 7 , for example, makes the linear regression

TABLE 4-6

Two Coefficients of the Sample Linear Regression of $\hat{\tau}^*_V$ on τ, in Each of the Five Cases.

Case	α	β
4	0.99588	0.00111
5	0.99605	0.00057
6	0.99542	-0.00022
7	0.99673	0.00053
8	0.99599	0.00027

$0.98339\tau + 0.00043$, that for Case 6 makes it $0.96849\tau + 0.00563$, and that for Case 5 makes it $0.93922\tau + 0.01638$, all of which are flatter than the original linear regressions.

It is observed that, in all five cases, the least value of the finite maximum likelihood estimates which our hypothetical examinees obtained is -2.6518 , and the greatest value 2.7683 . These two values are larger in absolute values than $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$, respectively, for the four cases, Cases 4 through 7 , and only Case 8 provides us with $\hat{\tau}^*_{V\text{-min}}$ and $\hat{\tau}^*_{V\text{-max}}$ which are larger in absolute values than these two finite maximum likelihood estimates. This fact implies that, out of the five sets of intervals of τ and the corresponding pairs of alternative estimates, those of Case 8 may be the most suitable ones for Subtest 3 . This set of alternative estimates also gives us an approximate conditional unbiasedness of $\hat{\tau}^*_V$ for truncated intervals. Figure 4-8 presents the sample regression of $\hat{\tau}^*_V$ on τ in Case 8, for the truncated interval of τ , $(-2.430, 2.586)$, which is the same as the interval, $(\underline{\tau},\bar{\tau})$, in Case 4 . We can see in this figure that the sample regression still clusters around the unbiasedness line for this truncated interval. In contrast to this, Figure 4-9 presents the sample regression in Case 4 for the extended interval of τ , $(-3.0, 3.0)$. The awkward shapes of clusters around the two endpoints of the extended interval of τ indicates that the two alternative estimates in Case 4 fail to provide us with an approximate conditional unbiasedness of $\hat{\tau}^*_V$ for this extended interval of τ .

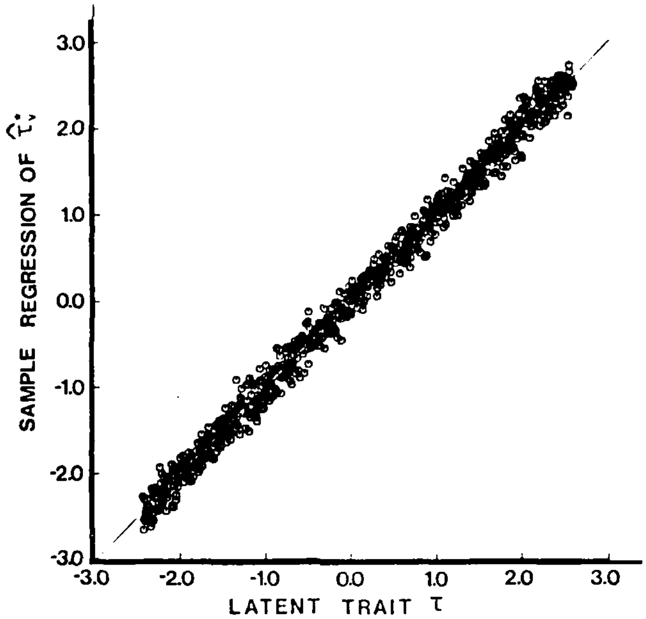


FIGURE 4-8

Sample Regression of $\hat{\tau}^*_V$ on τ : Case 8, Using the Interval (-2.430, 2.586), Instead of (-3.000, 3.000).

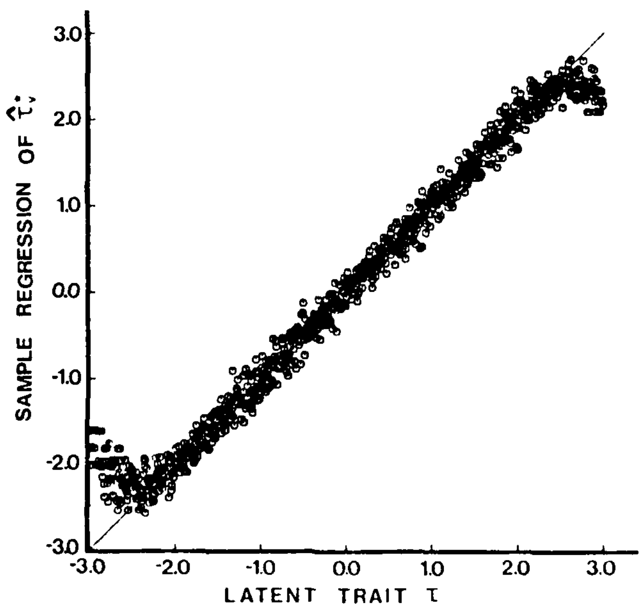


FIGURE 4-9

Sample Regression of $\hat{\tau}^*_V$ on τ : Case 4, Using the Interval (-3.000, 3.000), Instead of (-2.430, 2.586).

The error score for each individual examinee s is defined as was done in the preceding chapter, by replacing θ_s by τ_s, $\hat{\theta}^*_{V_s}$ by $\hat{\tau}^*_{V_s}$, and $I(\theta)$ by $I^*(\tau)$ in (3.17). For convenience, the same symbol, e_s, will be used for the error score defined for $\hat{\tau}^*_V$. Note that, in the present situation, the test information function, $I^*(\tau)$, is constant $(= 3.5^2)$ for the interval of τ of our interest, instead of being a unimodal function. The error score e_s was computed for each of the 4,180 examinees of Case 4, and each of the 5,000 examinees of Case 8. The frequency distributions of these error scores are presented as histograms in Figures 4-10 and 4-11, respectively, with the category width of 0.2, together with the standard normal density function. We can see that these two histograms are much closer to the standard normal density function, in comparison with those obtained in the preceding chapter upon the five hundred hypothetical examinees. It is also noted that these two resultant histograms are substantially different from each other, in spite of the fact that 4,125 error scores are common for these two histograms. The chi-square test for the goodness of fit of these two frequency distributions against the standard normal density function gives $\chi^2_0 = 23.3491$ with 29 degrees of freedom $(.70 < p < .80)$ for Case 4, and $\chi^2_0 = 55.6856$ with the same number of degrees of freedom $(.001 < p < .01)$ for Case 8. The mean, variance and standard deviation of the error score, e_s, are 0.0028, 1.0070 and 1.0035 for Case 4, and 0.0009, 0.9206 and 0.9595 for Case 8, respectively. It is interesting to note that

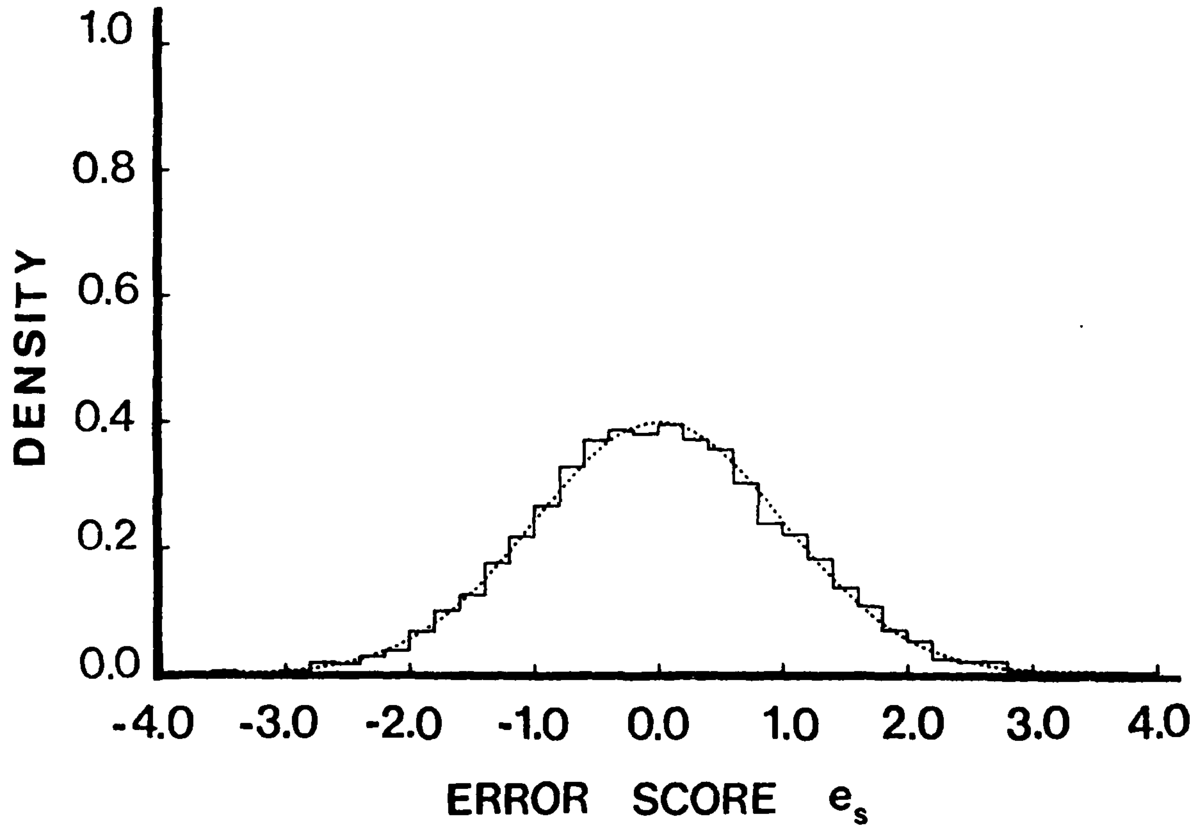


FIGURE 4-10

Frequency Distribution of the Error Score, e_s *, Which Is Based upon* Subtest 3, for the 4,180 Hypothetical Examinees of Case 4, Compared with the Standard Normal Density Function.

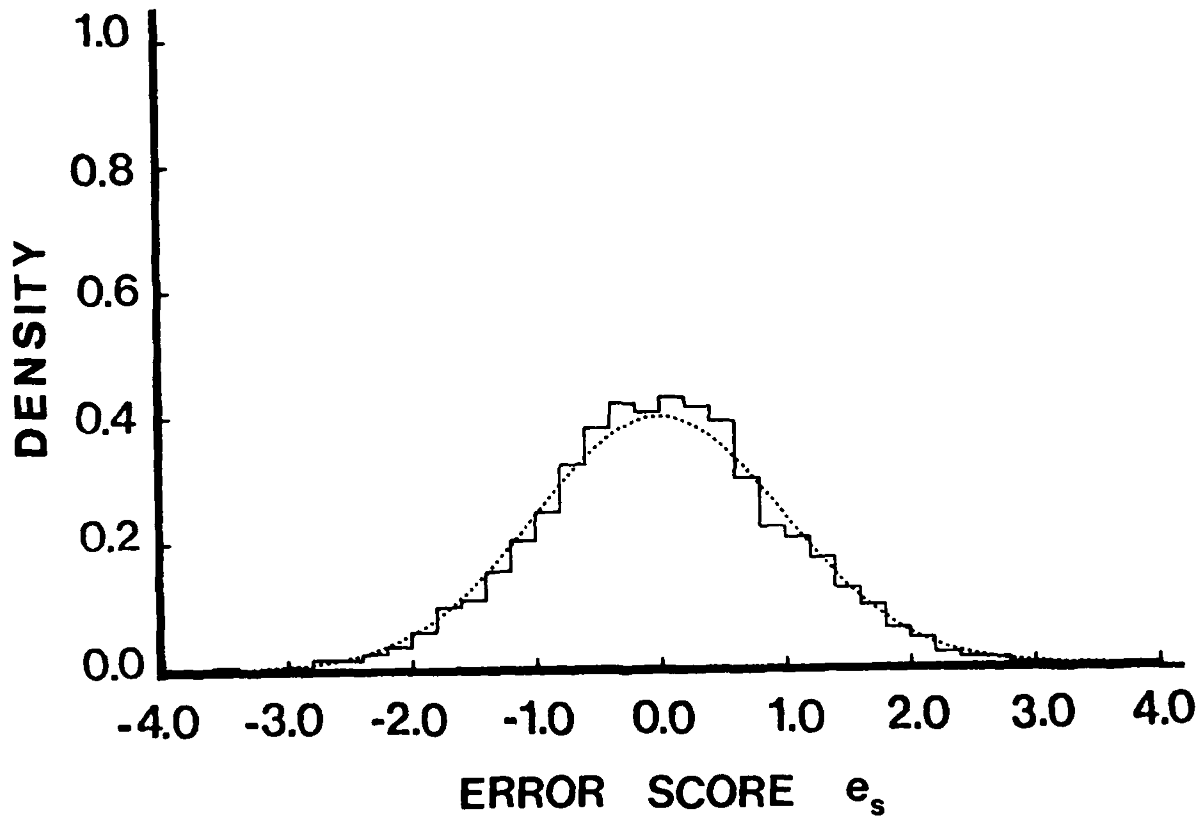


FIGURE 4-11

Frequency Distribution of the Error Score, e_s, Which Is Based upon Subtest 3, for the 5,000 Hypothetical Examinees of Case 8, Compared with the Standard Normal Density Function.

the dispersion of the error score in Case 8 is substantially less than unity. The correlation coefficient between τ and $\hat{\tau}^{*\prime}_{V}$ is -0.021 for Case 4 , and -0.025 for Case 8 . The sample linear regression of the error score e_s on τ is $-0.01444\tau + 0.00387$ in Case 4 , and $-0.01403\tau + 0.00093$ in Case 8 , both of which are very close to zero.

The pair of estimates, $\mu^{*\prime}_{1V\text{-}min}$ and $\mu^{*\prime}_{1V\text{-}max}$, which were introduced in the preceding chapter, were also obtained with respect to τ , for each of Cases 1 through 8 . These results are presented in Table 4-7 . We can see that for larger intervals of τ , like those in Cases 6 through 8 , the resultant estimates are similar to the corresponding values of $\hat{\tau}^{*}_{V\text{-}min}$ and $\hat{\tau}^{*}_{V\text{-}max}$, respectively.

Table 4-7

Bayes Estimates with the Uniform Prior for the Two Extreme Response Patterns, $\mu^{*\prime}_{1V\text{-}min}$ and $\mu^{*\prime}_{1V\text{-}max}$, Obtained with Respect to τ, for Each of the Eight Cases.

Case	$\mu^{*\prime}_{1V\text{-}min}$	$\mu^{*\prime}_{1V\text{-}max}$
1	-1.7434	1.9980
2	-1.9286	2.1810
3	-2.0965	2.3457
4	-2.2464	2.4905
5	-2.3143	2.5551
6	-2.4364	2.6684
7	-2.5402	2.7171
8	-2.7527	2.7805

V Discussion and Conclusions

The modified maximum likelihood estimate, θ^*_V , and its variation, $\hat{\theta}^*_V$, have been introduced and investigated, in comparison with Bayesian estimates and another population-free estimate, $\mu^{*\prime}_{1V}$. The former of these two newly proposed estimates is effective when a given test is short, like LIS-U , and the latter is useful when it is longer and more informative, like Subtest 3 .

The basic idea behind this research is to admit that each test has a certain limited range of ability for which it is effective in estimating the examinee's ability. Although this is a self-evident fact, for some reason, the idea has not fully been accepted by many researchers, and people tend to use tests for overly wide ranges of ability, and turn to inappropriate methods like Bayesian estimation, in order to make the result plausible. This is evidently a false solution, using the pretense that the test has measured something, while it has failed in so doing, and it is mainly an arbitrarily set prior which has given the examinee his ability score. The greatest fault of the Bayesian estimation may be that it is against the principle of objective testing, since it contaminates the resulting ability estimate by something other than the examinee's performance in the test.

The conditional unbiasedness of the ability estimate, given ability, is by far the most important in order to sustain the principle of the objectivity of testing. Taking this fact in mind, we can still try to enhance the usefulness of a given test, by expanding the range

of ability for which the test is effective. One way of doing this is to provide a suitable estimator.

The maximum likelihood estimator has a useful characteristic of asymptotic conditional unbiasedness. For less informative tests, however, the conditional probability with which the examinee obtains one of the two extreme response patterns, V-min and V-max , given ability, is substantially high, and this asymptotic characteristic cannot be used as an approximation. Thus, in such a case, we will see extreme values like negative and positive infinities among the maximum likelihood estimates of our examinees: the fact that restricts the effectiveness of the test.

The two modified maximum likelihood estimates, θ^*_V and $\hat{\theta}^*_V$, which were proposed and discussed in the present paper, were conceived with the following considerations in mind.

(1) We follow the principle of the objectivity of testing, and, in estimating his ability, we use nothing but the examinee's performance on the test.

(2) The resultant estimate provides us with an approximate conditional unbiasedness of estimation.

(3) The range of ability for which the test is effective is enhanced.

This has been done by replacing the maximum likelihood estimates for the two extreme response patterns, V-min and V-max , by $\theta^*_{V\text{-min}}$ and $\theta^*_{V\text{-max}}$, or by $\hat{\theta}^*_{V\text{-min}}$ and $\hat{\theta}^*_{V\text{-max}}$, respectively. The results

proved to be promising.

One distinct feature of the present study may be the use of the Monte Carlo method in obtaining $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$. One may argue that, because of this, we cannot avoid the sampling fluctuations of $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$. While it is true, this can be minimized by using a large enough sample size. In the present study, we used as large a sample size as 5,000 , and this can be made even larger, if we wish. In any case, even if we use, say, ten thousand hypothetical examinees, it is still a reduction, considering that even a relatively short test like Subtest 3 contains as many as 14,348,907 different response patterns.

An alternative way may be the use of the estimates, $\mu^{*\prime}_{1V\text{-}min}$ and $\mu^{*\prime}_{1V\text{-}max}$. Since these estimates do not depend upon the Monte Carlo method, or any samples, they do not have the problem of sampling fluctuations. The approximate conditional unbiasedness may not be reached just as well, however, if we use $\mu^{*\prime}_{1V\text{-}min}$ and $\mu^{*\prime}_{1V\text{-}max}$ instead of $\hat{\theta}^*_{V\text{-}min}$ and $\hat{\theta}^*_{V\text{-}max}$.

REFERENCES

[1] Elderton, W. P. and N. L. Johnson. Systems of frequency curves. Cambridge University Press, 1969.

[2] Indow, T. & Samejima, F. LIS measurement scale for non-verbal reasoning ability. Tokyo: Nippon Bunka Kagakusha, 1962. (in Japanese)

[3] Indow, T. & Samejima, F. On the results obtained by absolute scaling model and the Lord model in the field of intelligence. Yokohama: Psychological Laboratory, Hiyoshi Campus, Keio University, 1966.

[4] Johnson, N. L. & S. Kotz. Continuous univariate distributions - I and II. Houghton Mifflin, 1970.

[5] Lord, F. M. & M. R. Novick. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

[6] Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 1969.

[7] Samejima, F. A general model for free-response data. Psychometrika Monograph, No. 18, 1972.

[8] Samejima, F. Graded response model of the latent trait theory and tailored testing. Proceedings of the First Conference on Computerized Adaptive Testing, 1975, Civil Service Commission and Office of Naval Research, 1975, pages 5-17.

[9] Samejima, F. Effects of individual optimization in setting the boundaries of dichotomous items on accuracy of estimation. Applied Psychological Measurement, 1977a, 1, 77-94.

[10] Samejima, F. A use of the information function in tailored testing. Applied Psychological Measurement, 1977b, 1, 233-247.

[11] Samejima, F. A method of estimating item characteristic functions using the maximum likelihood estimate of ability. Psychometrika, 42, 1977c, pages 163-191.

LIST OF ONR TECHNICAL REPORTS

(1) Samejima, F. Estimation of the operating characteristics of item response categories I: Introduction to the Two-Parameter Beta Method. Office of Naval Research, RR-77-1, 1977.

(2) Samejima, F. Estimation of the operating characteristics of item response categories II: Further development of the Two-Parameter Beta Method. Office of Naval Research, RR-78-1, 1978.

(3) Samejima, F. Estimation of the operating characteristics of item response categories III: The Normal Approach Method and the Pearson System Method. Office of Naval Research, RR-78-2, 1978.

(4) Samejima, F. Estimation of the operating characteristics of item response categories IV: Comparison of the different methods. Office of Naval Research, RR-78-3, 1978.

(5) Samejima, F. Estimation of the operating characteristics of item response categories V: Weighted Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, RR-78-4, 1978.

(6) Samejima, F. Estimation of the operating characteristics of item response categories VI: Proportioned Sum Procedure in the Conditional P.D.F. Approach. Office of Naval Research, RR-78-5, 1978.

(7) Samejima, F. Estimation of the operating characteristics of item response categories VII: Bivariate P.D.F. Approach with Normal Approach Method. Office of Naval Research, RR-78-6, 1978.

(8) Samejima, F. Constant Information Model: A new, promising item characteristic function. Office of Naval Research, RR-79-1, 1979.

(9) Samejima, F. and P. S. Livingston. Method of moments as the least squares solution for fitting a polynomial. Office of Naval Research, RR-79-2, 1979.

(10) Samejima, F. Convergence of the conditional distribution of the maximum likelihood estimate, given latent trait, to the asymptotic normality: Observations made through the constant information model. Office of Naval Research, RR-79-3, 1979.

(11) Samejima, F. A new family of models for the multiple-choice item. Office of Naval Research, RR-79-4, 1979.

(12) Samejima, F. Research on the multiple-choice test item in Japan: Toward the validation of mathematical models. Office of Naval Research, Tokyo, Scientific Monograph 3, April, 1980.

(13) Samejima, F. & R. L. Trestman. Analysis of Iowa data I: Initial study and findings. Office of Naval Research , RR-80-1, 1980.

(14) Samejima, F. Estimation of the operating characteristics when the test information of the old test is not constant I: Rationale. Office of Naval Research, RR-80-2, 1980.

(15) Samejima, F. Is Bayesian estimation proper for estimating the individual's ability? Office of Naval Research, RR-80-3, 1980.

(16) Samejima, F. Estimation of the operating characteristics when the test information of the old test is not constant II: Simple sum procedure of the conditional P.D.F. approach/normal approach method using three subtests of the old test. Office of Naval Research, RR-80-4, 1980.

AD-A104 088 TENNESSEE UNIV KNOXVILLE DEPT OF PSYCHOLOGY F/G 12/1
AN ALTERNATIVE ESTIMATOR FOR THE MAXIMUM LIKELIHOOD ESTIMATOR F--ETC(U)
JUN 81 F SAMEJIMA N00014-77-C-0360
UNCLASSIFIED RR-81-1 NL

2 OF 2
AD
A104 088

END
DATE
FILMED
10-81
DTIC

Blank

APPENDIX

TABLE A-1

Comparison of Two Population-Free Estimates, θ^*_V and $\mu^{*\prime}_{1V}$, and Two Bayesian Estimates, $\mu^{\prime}_{1V}$ and $\hat{\theta}_V$, for Each of the One Hundred and Twenty-Eight Response Patterns of LIS-U .

	V	θ^*_V	$\mu^{*\prime}_{1V}$	$\mu^{\prime}_{1V}$	$\hat{\theta}_V$
1	1111111	-1.9254	-1.6515	-1.3764	-1.3617
2	2111111	-1.2260	-1.2534	-0.9328	-0.9267
3	1211111	-0.8567	-0.9823	-0.6905	-0.7017
4	2211111	-0.4968	-0.5126	-0.4146	-0.4104
5	1121111	-1.1661	-1.2069	-0.9098	-0.9067
6	2121111	-0.7267	-0.7863	-0.6065	-0.5943
7	1221111	-0.5030	-0.5175	-0.4237	-0.4198
8	2221111	-0.1602	-0.1619	-0.1549	-0.1331
9	1112111	-1.3167	-1.3188	-0.9942	-0.9878
10	2112111	-0.8129	-0.8821	-0.6696	-0.6573
11	1212111	-0.5726	-0.5929	-0.4790	-0.4766
12	2212111	-0.2307	-0.2383	-0.2115	-0.1920
13	1122111	-0.7983	-0.8610	-0.6655	-0.6544
14	2122111	-0.4491	-0.4914	-0.3994	-0.3765
15	1222111	-0.2520	-0.2591	-0.2295	-0.2117
16	2222111	0.0997	0.1056	0.0442	0.0828
17	1111211	-1.0350	-1.0910	-0.8358	-0.8325
18	2111211	-0.6555	-0.7047	-0.5577	-0.5450
19	1211211	-0.4522	-0.4627	-0.3868	-0.3810
20	2211211	-0.1291	-0.1237	-0.1273	-0.1047
21	1121211	-0.6523	-0.6966	-0.5595	-0.5476
22	2121211	-0.3400	-0.3683	-0.3128	-0.2888
23	1221211	-0.1515	-0.1506	-0.1487	-0.1280
24	2221211	0.2005	0.2150	0.1247	0.1666
25	1112211	-0.7231	-0.7758	-0.6145	-0.6030
26	2112211	-0.4021	-0.4364	-0.3636	-0.3406
27	1212211	-0.2151	-0.2149	-0.2008	-0.1829
28	2212211	0.1241	0.1325	0.0652	0.1037
29	1122211	-0.4125	-0.4438	-0.3734	-0.3529
30	2122211	-0.1128	-0.1374	-0.1365	-0.1018
31	1222211	0.0832	0.0906	0.0353	0.0702
32	2222211	0.4735	0.4949	0.3279	0.3890
33	1111121	-0.7831	-0.8271	-0.6681	-0.6621
34	2111121	-0.4882	-0.5161	-0.4326	-0.4179
35	1211121	-0.3170	-0.3178	-0.2816	-0.2718
36	2211121	-0.0147	-0.0035	-0.0372	-0.0124
37	1121121	-0.4937	-0.5133	-0.4394	-0.4254
38	2121121	-0.2223	-0.2364	-0.2168	-0.1918
39	1221121	-0.0448	-0.0352	-0.0612	-0.0393
40	.2221121	0.2950	0.3193	0.2042	0.2467
41	1112121	-0.5503	-0.5814	-0.4863	-0.4729
42	2112121	-0.2767	-0.2952	-0.2627	-0.2390
43	1212121	-0.1045	-0.0993	-0.1107	-0.0896
44	2212121	0.2192	0.2374	0.1453	0.1847
45	1122121	-0.2915	-0.3063	-0.2756	-0.2533
46	2122121	-0.0260	-0.0317	-0.0553	-0.0209
47	1222121	0.175[illegible]	0.1914	0.1131	0.1490
48	2222121	[illegible]	[illegible]	[illegible]	0.4641
49	1111221	[illegible]	-0.47[illegible]	-0.407[illegible]	-0.3927
50	2111221	-0.192[illegible]	-0.20[illegible]	-0.192[illegible]	-0.1671
51	1211221	-0.0196	-0.0084	-0.0402	-0.0168
52	2211221	0.3116	0.3378	0.2196	0.2619
53	1121221	-0.2102	-0.2258	-0.207[illegible]	-0.1827
54	2121221	0.0524	0.0505	0.0095	0.0455
55	1221221	0.2618	0.2852	0.1839	0.2224
56	2221221	0.6773	0.7224	0.4871	0.5518
57	1112221	-0.2635	-0.2744	-0.2501	-0.2275
58	2112221	-0.0318	-0.0072	-0.0359	-0.0016
59	1212221	0.1938	0.2114	0.1297	0.1656
60	2212221	0.5722	0.6086	0.4144	0.4742

TABLE A-1 (Continued)

	V	θ^*_V	$\mu^{*\prime}_{1V}$	μ'_{1V}	$\hat{\hat{\theta}}_V$
61	1122221	-0.0275	-0.0329	-0.0564	-0.0241
62	2122221	0.2443	0.2481	0.1661	0.2123
63	1222221	0.5030	0.5353	0.3679	0.4218
64	2222221	1.0301	1.0679	0.7218	0.8086
65	1111112	-0.7214	-0.7651	-0.6251	-0.6151
66	2111112	-0.4411	-0.4696	-0.3935	-0.3803
67	1211112	-0.2772	-0.2755	-0.2527	-0.2391
68	2211112	0.0138	0.0226	-0.0150	0.0118
69	1121112	-0.4485	-0.4738	-0.4063	-0.3890
70	2121112	-0.1881	-0.2025	-0.1909	-0.1636
71	1221112	-0.0164	-0.0088	-0.0399	-0.0141
72	2221112	0.3085	0.3296	0.2161	0.2604
73	1112112	-0.5022	-0.5326	-0.4512	-0.4345
74	2112112	-0.2404	-0.2587	-0.2352	-0.2090
75	1212112	-0.0739	-0.0703	-0.0873	-0.0637
76	2212112	0.2364	0.2516	0.1595	0.2008
77	1122112	-0.2562	-0.2733	-0.2487	-0.2239
78	2122112	0.0008	-0.0071	-0.0354	0.0007
79	1222112	0.1936	0.2073	0.1250	0.1659
80	2222112	0.5608	0.5895	0.4060	0.4672
81	1111212	-0.4110	-0.4319	-0.3763	-0.3588
82	2111212	-0.1609	-0.1723	-0.1680	-0.1406
83	1211212	0.0058	0.0160	-0.0202	0.0059
84	2211212	0.3238	0.3462	0.2305	0.2745
85	1121212	-0.1794	-0.1902	-0.1834	-0.1576
86	2121212	0.0734	0.0712	0.0268	0.0643
87	1221212	0.2759	0.2960	0.1958	0.2361
88	2221212	0.6651	0.7049	0.4854	0.5501
89	1112212	-0.2277	-0.2417	-0.2248	-0.2000
90	2112212	0.0213	0.0155	-0.0172	0.0187
91	1212212	0.2108	0.2258	0.1436	0.1813
92	2212212	0.5698	0.6001	0.4166	0.4767
93	1122212	-0.0043	-0.0099	-0.0376	-0.0038
94	2122212	0.2571	0.2599	0.1779	0.2247
95	1222212	0.5042	0.5319	0.3727	0.4267
96	2222212	0.9851	1.0232	0.7253	0.7902
97	1111122	-0.3069	-0.3162	-0.2900	-0.2713
98	2111122	-0.0755	-0.0790	-0.0942	-0.0664
99	1211122	0.0996	0.1051	0.0507	0.0779
100	2211122	0.4029	0.4340	0.2983	0.3426
101	1121122	-0.0958	-0.0995	-0.1112	-0.0849
102	2121122	0.1455	0.1497	0.0908	0.1281
103	1221122	0.3518	0.3797	0.2614	0.3021
104	2221122	0.7529	0.8052	0.5556	0.6195
105	1112122	-0.1410	-0.1473	-0.1504	-0.1250
106	2112122	0.0950	0.0961	0.0481	0.0838
107	1212122	0.2863	0.3083	0.2091	0.2474
108	2212122	0.6501	0.6918	0.4837	0.5434
109	1122122	0.0679	0.0686	0.0264	0.0602
110	2122122	0.3222	0.3315	0.2365	0.2828
111	1222122	0.5795	0.6173	0.4363	0.4906
112	2222122	1.0935	1.1411	0.7851	0.8674
113	1111222	-0.0751	-0.0771	-0.0952	-0.0668
114	2111222	0.1558	0.1662	0.1049	0.1420
115	1211222	0.3644	0.3930	0.2734	0.3139
116	2211222	0.7577	0.8117	0.5632	0.6259
117	1121222	0.1313	0.1360	0.0814	0.1165
118	2121222	0.3706	0.4052	0.2945	0.3424
119	1221222	0.6758	[illegible]	0.5121	0.5670
120	2221222	1.3028	[illegible]	0.9009	0.9840
121	1112222	0.0842	0.0852	0.0412	0.0748
122	2112222	0.3334	0.3433	0.2477	0.2935
123	1212222	0.5873	0.6251	0.4453	0.4589
124	2212222	1.0967	1.1433	0.7396	0.8695
125	1122222	0.2963	0.3097	0.2191	0.2622
126	2122222	0.5807	0.6035	0.4476	0.5055
127	1222222	0.9714	1.0291	0.7175	0.7858
128	2222222	1.6473	1.6430	1.2695	1.3542

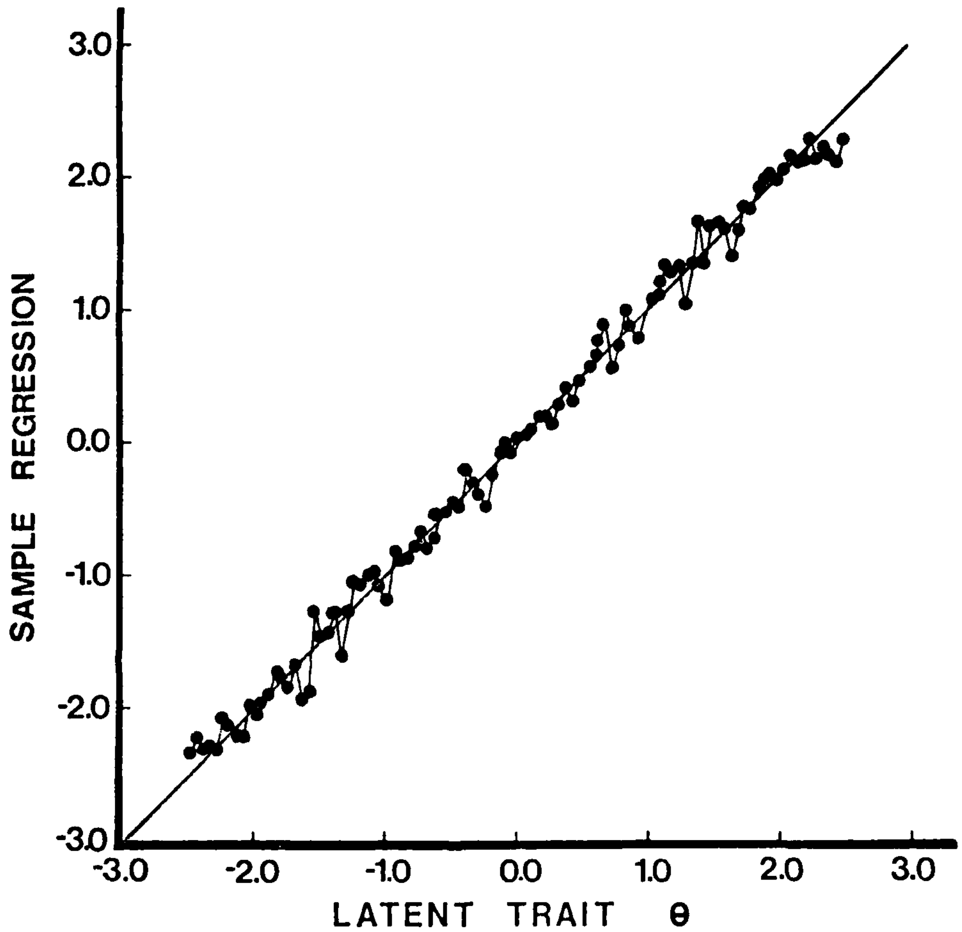


FIGURE A-1

Sample Regression of the Modified Maximum Likelihood Estimate, $\hat{\theta}^*_V$, for the One Hundred Ability Levels : Each Point is the Mean of Five Values of $\hat{\theta}^*_V$.

DISTRIBUTION LIST

Navy

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 COMNAVMILPERSCOM (N-6C)
Dept. of Navy
Washington, DC 20370

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Henry M. Halff
Department of Psychology,C-009
University of California at San Diego
La Jolla, CA 92093

1 Dr. Patrick R. Harrison
Psychology Course Director
LEADERSHIP & LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

Navy

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. George Moeller
Head, Human Factors Dept.
Naval Submarine Medical Research Lab
Groton, CN 06340

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

Navy

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
(OP-115)
Washington, DC 20350

1 Dr. Donald F. Parker
Graduate School of Business Administrati
University of Michigan
Ann Arbor, MI 48109

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 Director, Research & Analysis Division
Plans and Policy Department
Navy Recruiting Command
4015 Wilson Boulevard
Arlington, VA 22203

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

Navy

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr John H. Wolfe
Code P310
U. S. Navy Personnel Research and
Development Center
San Diego, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Dexter Fletcher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria,VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milt Maier
U.S. Army Reasearch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 LTC Michael Plummer
Chief, Leadership & Organizational
Effectiveness Division
Office of the Deputy Chief of Staff
for Personnel
Dept. of the Army
Pentagon, Washington DC 20301

1 DR. JAMES L. RANEY
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

Army

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commandant
US Army Institute of Administration
Attn: Dr. Sherrill
FT Benjamin Harrison, IN 46256

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

1 U.S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Research and Measurment Division
Research Branch, AFMPC/MPCYPR
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway
Technical Director
AFHRL(OT)
Williams AFB, AZ 58224

1 Dr. Frank Schufletowski
U.S. Air Force
ATC/XPTD
Randolph AFB, TX 78148

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Major Howard Langdon
Headquarters, Marine Corps
OTTI 31
Arlington Annex
Columbia Pike at Arlington Ridge Rd.
Arlington, VA 20380

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 Major Michael L. Patrow, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 Chief, Psychological Reserch Branch
U. S. Coast Guard (G-P-1/2/TP42)
Washington, DC 20593

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Dr. William Graham
Testing Directorate
MEPCOM/MEPCT-P
Ft. Sheridan, IL 60037

1 Director, Research and Data
OASD(MRA&L)
3B919, The Pentagon
Washington, DC 20301

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

1 DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civil Govt

1 Dr. Lorraine D. Eyde
Personnel R&D Center
Office of Personnel Management of USA
1900 EStreet NW
Washington, D.C. 20415

1 Jerry Lehnus
REGIONAL PSYCHOLOGIST
U.S. Office of Personnel Management
230 S. DEARBORN STREET
CHICAGO, IL 60604

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Jackson Beatty
Department of Psychology
University of California
Los Angeles, CA 90024

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. John Bergan
School of Education
University of Arizona
Tuscon AZ 85721

1 Dr. Werner Birke
DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. R. Darrel Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

Non Govt

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. Anthony Cancelli
School of Education
University of Arizona
Tuscon, AZ 85721

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William E. Coffman
Director, Iowa Testing Programs
334 Lindquist Center
University of Iowa
Iowa City, IA 52242

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

Non Govt

1 Director
Behavioural Sciences Division
Defence & Civil Institute of
Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

Non Govt

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Frederick Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Jack Hunter
2122 Coolidge St.
Lansing, MI 48906

Non Govt

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

Non Govt

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Professor Jason Millman
Department of Education
Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

Non Govt

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

Non Govt

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 David E. Stone, Ph.D.
Hazeltine Corporation
7680 Old Springhouse Road
McLean, VA 22102

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Psychometric Research Group
Educational Testing Service
Princeton, NJ 08541

1 Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2

Non Govt

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

Navy

1 Mr. Donald Calder
Office of Naval Research
325 Hinman Research Building
Atlanta, GA 30332

Army

1 Dr. Randall M. Chambers
U.S. Army Research Institute
for the Behavioral & Social
Sciences
Fort Sill Field Unit
P.O. Box 3066
Fort Sill, OK 73503

Non Govt

1 Dr. Bert F. Green
Department of Psychology
The John's Hopkins University
Charles at 34th Street
Baltimore, MD 21218

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Dr. William W. Turnbull
Educational Testing Service
Princeton, NJ 08540

1 Dr. Isaac I. Bejar
Department of Psychology
Elliot Hall
75 East River Road
Minneapolis, MN 55455

1 Dr. George Woods
1106 Newport Ave.
Victoria, B. C.
V8S 5E4 Canada

1 Dr. Lowell Schipper
Department of Psychology
Bowling Green State University
Bowling Green, OH 43403

Non Govt

1 Dr. P. Mengal
Faculte' de Psychologie
et des Sciences de l'Education
Universite' de Geneve
3 fl. de l'Universite
1201 Geneva SWITZERLAND

1 Dr. Wim J. van der Linden
Vakgroep Onderwisjskunde
Postbus 217
7500 EA Enschede
The Netherlands

1 Dr. Lutz Hornke
University Duesseldorf
Erz. Wiss.
D-4000 Duesseldorf
WEST GERMANY

1 Dr. Wolfgang Buchtala
8346 Simbach Inn
Postfach 1306
Industriestrasse 1
WEST GERMANY

1 Dr. Sukeyori Shiba
Faculty of Education
University of Tokyo
Hongo, Bumkyoku
Tokyo, Japan 113

1 Mr. Yukihiro Noguchi
Faculty of Education
University of Tokyo
Hongo, Bumkyoku
Tokyo, Japan 113

1 Dr. Takahiro Sato (Representative)
Application Research Laboratory
Central Research Laboratories
Nippon Electric Co., Ltd.
4-1-1 Miyazaki, Takatsu-ku
Kawasaki 213, Japan

Non Govt

1 Dr. James Chen
354 Lindquist Center for Measurement
University of Iowa
Iowa City, Iowa 52240

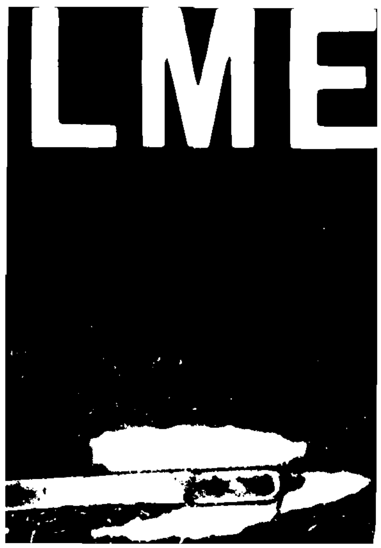
1 Mr. Jeffrey Jankowitz
Department of Educational Psychology
210 Education Building
University of Illinois
Urbana, IL 61801

1 Dr. Douglas Carroll
Bell Laboratories
600 Mountain Ave.
Murray Hill, N.J. 07974

1 Dr. Robert Guion
Department of Psychology
Bowling Green State University
Bowling Green, OH 43403

1 Dr. Niel Timm
Department of Information Systems
Planning
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Albert Beaton
Educational Testing Service
Princeton, New Jersey 08450


LME